DOCUMENT RESUME

ED 445 067 TM 031 683

AUTHOR	Fiske, Edward B.
TITLE	Education for All: Status and Trends, 2000. Assessing Learning Achievement.
INSTITUTION	United Nations Educational, Scientific, and Cultural Organization, Paris (France). Education for All Forum Secretariat.
ISSN	ISSN-1020-0908
PUB DATE	2000-00-00
NOTE	76p.; For an earlier issue, see ED 380 856. Photographs may not reproduce clearly.
AVAILABLE FROM	For full text (English, French, and Spanish): http://www.education.unesco.org/efa.
PUB TYPE	Reports - Evaluative (142)
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	*Academic Achievement; Educational Research; *Educational Trends; Elementary Secondary Education; Foreign Countries; *Measurement Techniques; Outcomes of Education; *Status; *Student Evaluation
IDENTIFIERS	World Conference on Education for All

ABSTRACT

This fifth issue of "Education for All, Status and Trends," examines the reasons for the growing global interest in educational measurement, lays out some of the central findings of the movement, and then takes up the all-important question of how assessment of student outcomes can be put to the service of promoting quality education for all children. Research growing from the 1990 World Conference on Education for All has shown how the experience of developed and developing countries in carrying out national assessments has led to a body of general principles about how best to organize such measurement exercises and how to make the most effective use of the data they generate. Following an introduction, these principles are developed through the following sections: (1) "How Countries Measure Student Achievement"; (2) "How Measurement Can Improve Teaching and Learning"; (3) "Recent Research on the Quality of Education"; (4) "Broad Findings of Recent Research"; (5) "Factors That Contribute to Achievement"; (6) "How Countries Use Assessment To Improve Student Learning"; and (7) "Issues of Measurement." An appendix contains five statistical tables for specific world areas. (Contains 27 figures.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

Education for All

Status and Trends 2000

World Education Forum
26-28 April 2000
Dakar Senegal

TM

ED 445 067

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
F. Zanuttini
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
1

TM031683

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Assessing learning achievement

INTERNATIONAL CONSULTATIVE FORUM ON EDUCATION FOR ALL

BEST COPY AVAILABLE

Education for All

Status and Trends 2000

World Education Forum
26-28 April 2000
Dakar Senegal

Assessing learning achievement

INTERNATIONAL CONSULTATIVE FORUM ON EDUCATION FOR ALL

BEST COPY AVAILABLE

This report is published by UNESCO for the INTERNATIONAL CONSULTATIVE FORUM ON EDUCATION FOR ALL, the global mechanism established to promote and monitor progress towards Education for All goals.
The designations employed and the presentation of material in this report do not imply the expression of any opinion whatsoever on the part of UNESCO concerning the legal status of any country, territory, city or area or of its authorities, or concerning its frontiers or boundaries. Any part of Education for All: Status and Trends may be freely reproduced with the appropriate acknowledgement.
For further information, please contact:

EFA Forum Secretariat
UNESCO
7 Place de Fontenoy
75352 Paris 07 SP
France
Tel: +33 (0)1 45 68 15 24
Fax: +33 (0)1 45 68 56 29
E-mail: efa@unesco.org
Internet web site: www.education.unesco.org/efa

Editorial coordination ▸ Ulrika PEPPLER BARRY

Principal author ▸ Edward B. FISKE

Statistics ▸ Hilaire MPUTU, S. K. CHU

Technicals advisers ▸ Emilie BARRIER, Juan CASASSUS, Vinayagum CHINAPAH, Chantal PACTEAU, Neville POSTLETHWAITE, Kenneth ROSS, Ernesto SCHIEFELBEIN

Graphic design ▸ Sylvaine BAEYENS

Printing ▸ GRAPHOPRINT

Cover photo: Primary school, Shanghai, China.
(Photo: ©Eric FRANCESCHI/Agence VU)

BEST COPY AVAILABLE

©UNESCO 2000
ISSN 1020-0908
Printed in France

Contents

Preface 4

Introduction 6

Part I How Countries Measure Student Achievement 12
- School-Based Assessments
- Public Examinations
- National Assessments
- International Assessments

Part II How Measurement Can Improve Teaching and Learning 18
- Teaching to the Test
- Alignment of Curricula and Instruction
- Informing National Education Policy
- Increasing Public Support for Education
- Limitations of Various Forms of Measurement

Part III Recent Research on the Quality of Education 24
1. Southern Africa Consortium for Measuring Educational Quality (SACMEQ)
2. Monitoring Learning Achievement (MLA)
3. Latin American Laboratory for Assessment of Educational Quality (Laboratorio)
4. Third International Mathematics and Science Study (TIMSS)
5. Programme for the Analysis of Educational Systems of the CONFEMEN Countries (PASEC)
6. Programme for International Student Assessment (PISA)

Part VI Broad Findings of Recent Research 30
- The Low Level of Achievement in Developing Countries
- Changes in Pupil Performance Since Jomtien
- Differences Within Countries Exceed Those Between Them
- Regional Disparities in Pupil Performance
- Achievement Differences Between Types of Schools and Groups of Pupils
- The Importance of Disaggregating Data

Part V Factors that Contribute to Achievement 48
- Relative Impact of Contextual and School-Related Factors
- Effects of Contextual Factors on Achievement
- Effects of School-Related Factors on Achievement

Part VI How Countries Use Assessment to Improve Student Learning 56
- Targeting of Scarce Resources
- Shaping Classroom Teaching
- Influencing Textbooks and Curricula
- Building Public Support for Education
- Assessing Learning Readiness

Part VII Issues of Measurement 60
- What to Assess
- Sampling
- Quality of Data
- Security
- Political influences

Conclusions 65

Annexes 68

BEST COPY AVAILABLE

Primary school [illegible] DESCAMPS/Agence

BEST COPY AVAILABLE

Preface

How much are our children learning at school? Is it enough? Are there major gaps in their learning that need to be addressed? Parents around the world ask these questions, and it is likely that many of them are relatively satisfied with the answers. They have reason to be confident that their children are acquiring the knowledge and skills that they will need to function as workers, family members and citizens in the 21st century.

For tens of millions of other parents, however, these questions are deeply troubling. They and their children live in developing countries or in marginalized areas of the industrialized world where schools are underfunded, teachers lack adequate training and student achievement may or may not reach the level of basic literacy and numeracy. Such parents have good reason to doubt that the basic learning needs of their children are being met.

A number of research projects launched in the wake of the World Conference on Education for All, which met in Jomtien, Thailand in 1990, have documented the extent to which knowledge, like other forms of wealth, seems to be unevenly distributed in our world today. These studies have been carried out on international, regional and national bases, and they have documented disparities in learning achievement not only between the various regions and countries of the world but within particular nations. Such studies have also become increasingly sophisticated from a technical point of view. They have provided detailed insights into areas of curricular content and skills development which now allow us to focus on how the information gained through tests, assessments and other devices can be used to enhance teaching and learning. The development of scales of achievement assists in the setting of pedagogical priorities and the allocation of scarce financial and human resources. There is also new understanding which emphasizes that learning is a responsibility that is shared by the student, the school and the home.

What, then, can we learn from this new research? That is the question posed in this fifth issue of *Education for All, Status and Trends*. It examines the reasons for the growing global interest in educational measurement, lays out some of the central findings of the movement and then takes up the all-important question of how assessment of student outcomes can be put to the service of promoting quality education for all children.

Speaking for the EFA Secretariat, I want to express our appreciation to Edward Fiske, the principal author, S.K. Chu and Hilaire Mputu, UNESCO Institute for Statistics, to the following technical advisers: Emilie Barrier, Juan Casassus, Vinayagum Chinapah, Chantal Pacteau, Neville Postlethwaite, Ernesto Schiefelbein and Kennet Ross, as well as to David Archer, Bougouma Ngom, Moncef Moalla and Clinton Robinson, for their kind assistance in preparing this report.

Svein Osttveit
Executive Secretary

BEST COPY AVAILABLE

How much are our children learning at school? From generation to generation in Tunisia. © Tadeusz ROLKE, Agence VU

BEST COPY AVAILABLE

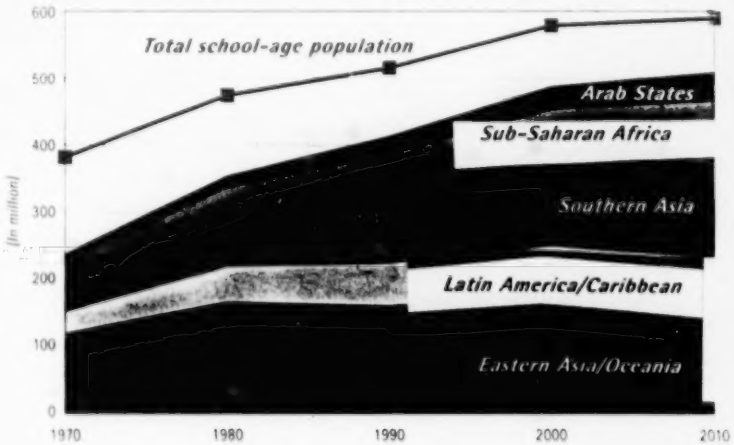
Introduction

Developed and developing countries alike understand that providing quality basic education for all children is essential not only to their own economic growth and social stability but to the functioning of a stable and equitable community of nations. Over the past decade a concentrated global effort has been made to increase the number of children in school. In 1990, the World Conference on Education for All, meeting in Jomtien, Thailand, urged all nations of the world to adopt policies that would ensure universal basic education by the year 2000.

Since Jomtien, considerable progress has been made in expanding the capacity of primary school systems in all regions of the world. Primary education in some developing countries has expanded to the extent that it reaches nearly all school-age children, and many of these countries have made significant efforts to overcome the gender gap in access to primary school. As shown in Figures 1 and 2, net enrolment has increased steadily in all developing regions, while the number of out-of-school children has decreased everywhere except in sub-Saharan Africa.

Figure 1

Primary-school age population: trends in enrolment in developing regions 1970–2010



Source: UNESCO Institute for Statistics. Estimates and projections of net enrolment ratios UNESCO, Paris, April 1999.

Although the rate of increase outpaced population growth, it has fallen far short of the pace necessary to meet the goal of universal enrolment by the year 2000. In the drive to provide universal primary education (UPE), the latest available information suggests that only the Eastern Asia/Oceania region is near that point, followed by the Latin America/Caribbean. Some progress can be observed in Southern Asia, but the number of out-of-school children is still high. In the Arab States, the non-enrolled school-age population has remained steady since 1970. Sub-Saharan Africa remains the only region where the non-enrolled school age population has increased since 1980, after a decrease between 1970 and 1980.

BEST COPY AVAILABLE

Figure 2

Primary-school age population: trends in numbers of non-enrolled children in developing regions 1970-2010

Source: UNESCO Institute for Statistics. Estimates and projections of net enrolment ratios, UNESCO, Paris, April 1999.

While giving all children the opportunity to attend school is obviously an important priority, it is but the first step towards the goal of 'education for all.' Once pupils find seats in a classroom, they need quality instruction; otherwise there will be little motivation to persist in school. In affirming the goal of universal basic education, participants in the Jomtien conference emphasized that reform efforts must focus on 'actual learning acquisition and outcome rather than exclusively upon enrolment.'[1] To drive home this point, participants urged countries to set specific qualitative targets. Learning achievement, they suggested, should be improved to the point that 'an agreed percentage of an appropriate age cohort – for example, 80 per cent of 14-year-olds – attains or surpasses a defined level of necessary achievement.'[2]

In order to talk seriously about educational quality, it is important to be able to define it and to measure it. Countries seeking to affirm the right of all children to a basic education need reliable means of describing the knowledge and skills that comprise such an education. They require the technical means and the organizational capacity to measure student achievement against these standards and to evaluate how well school systems are carrying out their own responsibilities.

They need expertise in translating assessment data into new instructional and governing policies that will increase the quality of teaching and learning.

Fortunately, the topic of educational measurement has received increased attention in recent years in developed and developing countries alike. Significant advances have been made in techniques for assessing educational achievement, including performance across national lines. Major research projects have been launched to investigate patterns of educational achievement and to explore both in-school and external factors that contribute to learning. Considerable thought has gone into ways in which developing countries can use data on student performance to make the most efficient and effective use of their educational resources.

The previous issue of *Status and Trends*[3] examined an important obstacle to effective teaching and learning at the primary level: the enormous wastage that results from high drop-out and repetition rates. This issue turns its attention to the bottom line of educational quality: actual student achievement as measured by tests and other assessment devices. Specifically, we will look at:

- What are current assessment practices in the world?
- What special research projects have recently been carried out to assess the quality of education in various countries.
- What have we learned about overall levels of achievement of primary-school pupils?
- What do we know about how achievement is distributed both between countries and between regions, types of schools and particular groups of pupils within the same country?
- What have assessment and research told us about factors that contribute to quality education?

We will then turn to the all-important policy question: *How assessment of student outcomes has been – and can be – used to promote quality education for all.*

1. World Declaration on Education for All: Meeting Basic Learning Needs, UNESCO, New York, 1990, para. 4.

2. Framework for Action to Meet Basic Learning Needs, Jomtien, Thailand, March 1990.

3. *Wasted Opportunities: When school fails*, Education for All: Status and Trends, UNESCO, Paris, 1998.

BEST COPY AVAILABLE

DEFINITION OF TERMS: ASSESSMENT, EVALUATION AND MEASUREMENT

Although the terms assessment, evaluation and measurement are often used interchangeably, it is important to note that these terms need to be interpreted separately and differently. Each frequently makes use of tests; nevertheless, none of these terms is synonymous with testing, and the types of tests required for each of the three processes may be very different. The three processes are considered below in reverse order.

Measurement

The regular dictionary definition of 'assigning a numerical quantity to' serves well in most applications of educational measurement. While instruments such as rulers and stopwatches can be used directly to measure height and speed, many characteristics of educational interest must be measured indirectly. Thus ability tests are typically used to measure such characteristics as intelligence and achievement tests are used to measure the amount of knowledge learned or forgotten. Measurement is not undertaken as an end in itself. It is a useful operation in the processes of evaluation, or for research where characteristics must be measured, or as part of the tasks of assessment of student performance.

Evaluation

In general, the use of the term 'evaluation' is reserved for application to abstract entities such as programmes, curricula and organizational situations. Its use implies a general weighing of the value or worth of something. Evaluation commonly involves making comparisons with a standard, or against criteria derived from stated objectives, or with other programmes, curricula or organizational situations. Evaluation is primarily an activity involved in research and development. It may require the measurement of educational outcomes, and it may involve the testing of both individuals and groups. Most judgements of an evaluative kind that are made in education would seem to be holistic in nature and to be based on a global examination of a situation.

Assessment

In general, the term 'assessment' should be reserved for use with reference to people. It may include administration of tests, or it may simply involve activities of grading or classifying according to some specified criteria. Student achievement in a particular course might be assessed, or students' attitudes towards particular aspects of their schooling might be examined. Such assessments are commonly based on an informal synthesis of a wide variety of evidence, although they might include the use of the test results, or responses to attitude scales and questionnaires.

Attention is increasingly being given to improving the quality of assessment, through the systematic specification of levels in performance assessment. However, such assessments can be converted to a scale of measurement through the use of the partial credit model, or the scaling of essay marks, and there is little need to view assessment as a process that does not involve measurement. It is, nevertheless, unfortunate that the term 'student evaluation' is now being widely used as a consequence of the growing emphasis on the evaluation of educational programmes and the financial support made available for such work.

Keeves, J.P. (Ed) Educational Research, Methodology, and Measurement: An International Handbook, 2nd ed. Pergamon, 1997, p. 711-12.

The growing interest in assessment

Increased attention to issues of educational measurement in recent years reflects a number of developments, including the following:

Focus on outcomes. In the past, political and educational leaders were generally content to evaluate schools and school systems on the basis of 'input' such as teacher preparation, the availability of textbooks and class size. The assumption was that such data could serve as a proxy for educational outcomes, which were difficult to measure. Substantial research, however, has now demonstrated the elusive nature between the various elements that go into schooling (especially those that are most easily measured) and student performance, and it has shown that the effectiveness of particular inputs varies widely from country to country. As a result, policy-makers have turned their attention from inputs to outputs, from process to results.

BEST COPY AVAILABLE

This new preoccupation with results has brought urgency to the task of specifying what they are. As Carol Bellamy, Executive Director of UNICEF put it, 'If the success of education is to be gauged by what and how children learn, better ways must be found to measure the quality and relevance of education.'[4]

Attention to accountability. Calls for greater accountability have been heard in all walks of life in recent years, especially in the delivery of public services such as health and education. These calls have been reinforced by the widespread trend toward the decentralization of large-scale enterprises, including state education systems. As local schools have been granted autonomy to manage their own affairs within nationally defined curricular and fiscal parameters, pressure has grown for measures that can be used to determine whether schools are carrying out their delegated responsibilities sufficiently well. Such pressures have in turn led to a broadened concept of accountability under which schools would be held accountable not only to government ministers but to the communities they serve and to civil society as a whole. Evaluation, which once focused almost exclusively on pupils and their performance, has been extended to schools, teachers and entire education systems.

Pressure to improve teaching and learning. In poor countries that have limited funds to invest in resources such as textbooks and teacher training, educators and economists alike are on the lookout for inexpensive means of improving the delivery of education. Policy-makers understand that tests and other forms of assessment can drive teaching and learning. Teachers tend to push subject matter that they know will be on examinations and assessments and to ignore other topics. Pupils work hardest to master material that they will encounter in testing situations, especially 'high-stakes' examinations that will directly affect their prospects for future schooling and jobs.

Policy-makers are also becoming increasingly attuned to the fact that some of the problems they face, such as the legacy of segregated schooling, may be unique to their particular country or to particular regions within it. Good information can inform strategic decisions relating to school construction, teacher training, curricula, textbooks and other instructional decisions. 'Examinations can be a powerful, low-cost means of influencing the quality of what teachers teach and what pupils learn in school,' wrote Stephen P. Heyneman and Angela W. Ransom. 'Examination agencies have an important role to play in increasing the effectiveness of schools.'[5]

Enrolment pressures. The advent of mass education at the primary level means that more pupils are seeking access to secondary and tertiary education, and, as a result, school systems face the need to make more and more decisions about which graduates will move on to higher levels of schooling. School-leaving examinations are widely viewed as a politically and ethically feasible way to make these decisions.[6]

Efficient allocation of resources. As they struggle to make the most of limited funds, a growing number of developing nations are coming to regard assessment as a source of guidance on how to invest scarce educational resources.

The median share of total government expenditures directed toward education in 1996 was 12.4 per cent in industrial countries and 16.3 per cent in developing nations. Given the magnitude of such investments, it makes sense to use every possible means to assure that education funds are spent both efficiently and effectively. ■

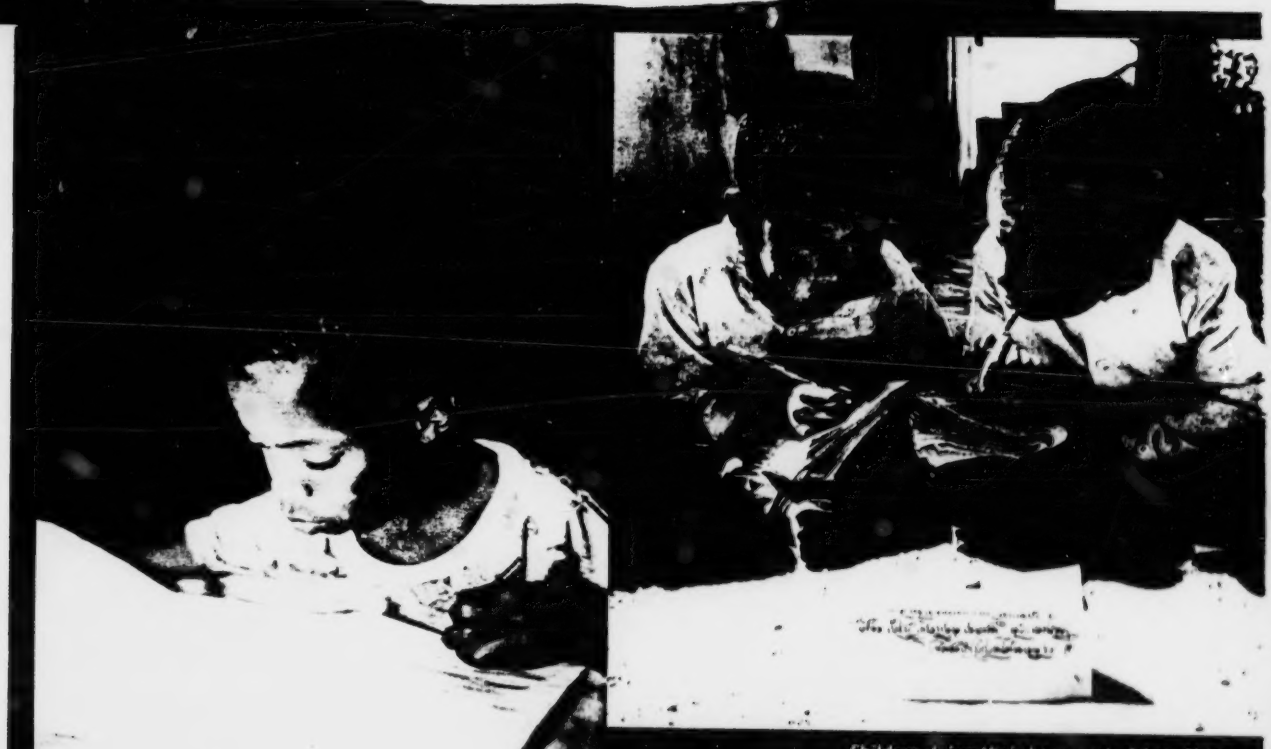
4. UNICEF, The State of the World's Children 1999.

5. Stephen P. Heyneman, and Angela W. Ransom, "Using Examinations and Testing to Improve Educational Quality," *Educational Policy*, Volume 4, Number 3, 1990.

6. Max A. Eckstein, and Harold J. Noah, *Comparing National Systems of Secondary School Leaving Examinations*, Pergamon Press.

BEST COPY AVAILABLE

Primary school, France. ©Gilles LARVOR/Agence VU



Children doing their homework, Tibet, China. ©Hien LAM DUC/Agence VU

Primary school, Dakar, Senegal. ©Hien LAM DUC/Agence VU

BEST COPY AVAILABLE

Primary school, Marseilles, France. © *Eric FRANCESCHI/Agence VU*

BEST COPY AVAILABLE

PART I

How countries measure student achievement

Broadly speaking, national education systems employ four types of measurement devices to monitor student achievement and schools performance at the primary level

School-based assessments

Many schools assess the progress of their pupils on a continuous basis by regularly measuring their performance against curricular goals and making the results available to classroom teachers. The results of such assessments can be used as a basis for part of the scoring of public examinations. Alternatively, they could become the sole basis for certification or selection.

School-based assessments are appealing because they offer immediate feedback to teachers on what pupils are learning and not learning, and thus become a basis for planning subsequent instruction. They are also more likely than standardized national tests to provide a full and accurate picture of what teachers are actually presenting to pupils. For example, such assessments can reflect practical topics that would not be anticipated by centralized examinations. They can also evaluate skills such as oral facility or the ability to organize a hands-on project that might not be captured by standardized tests.

BEST COPY AVAILABLE

Some developing countries, including Ethiopia and Lesotho, have built elements of school-based assessment into their public examination systems. Very few other countries, however, have followed their lead, and relatively little use is made of school-based assessment, which can impose considerable burdens on teachers and administrators.

Public examinations

Most countries, including many developing ones, have some sort of public examinations that individual pupils take at the end of primary school and other transition points in schooling. These examinations serve the purposes of:

1. selecting pupils to go on to secondary school,
2. certifying graduates for entry into the job market, and
3. fostering accountability for schools and school systems.

Of these, selection is usually the most important.

Characteristics of the public examination systems of developed countries are well known. The fiercely competitive Japanese system, which inspired the term 'examination hell,' has spawned a small industry of *juku*, or private cram schools, that propose to improve a candidate's chances of getting into a prestigious secondary school or university. For the last thirty years the University of Chile has administered an Aptitude Test that screens applicants to universities in that country. Secondary school pupils in France take *baccalauréat* examinations in order to qualify to enter universities, while their counterparts in Germany sit for examinations that lead to the awarding of the *abitur*, or 'exit credential,' that serves a similar purpose. In both countries the growth of pupils obtaining the secondary school-leaving credential has led to a progressive erosion of its capacity to assure entry into the most sought after universities. Such universities now have their own examinations and other entrance requirements.[7]

The United States, Canada and Sweden are the only major developed countries that do not have national public examinations. The United States has no national curriculum, and while the number has grown in recent years, only a minority – albeit a growing minority – of states sponsor tests to validate the awarding of a secondary school diploma. In the absence of a state-sponsored system, two private organizations sponsor examinations that American pupils take as part of the university admissions process.

Public examinations play an even more important role in developing countries, in Africa, Asia and the Caribbean, if only because alternative opportunities for advancement in such countries tend to be more limited.[8] Virtually all African countries conduct examinations at the end of primary, lower-secondary and upper-secondary cycles, and countries in French-speaking areas frequently require an additional exercise, such as end-of-year tests and competitive examinations for entry to subsequent levels of schooling. While most public examinations serve a certification function, the primary purpose is selection of pupils to move to the next level of the educational ziggurat. In Togo, for example, pupils take a highly competitive concours examination in order to proceed from the first to the second secondary cycle and again at the end of secondary schooling to move on to university or foreign study.[9]

China, which invented the Imperial Examination System in the tenth century to assure that the 'all-important business of government must not be left to the accidents of either birth or wealth,'[10] is another country that uses public examinations to rationalize the distribution of scarce places. Pupils take examinations at the end of the nine years of compulsory schooling in order to qualify for the various forms of upper secondary education. Subsequent exams at the end of upper secondary schooling determine university entrance.[11]

National assessments

In addition to public examinations established for purposes of selection and certification, many countries carry out regular and systematic measurement exercises designed to determine what students have learned as a result of their educational experiences. These are called 'national assessments.'

Assessments differ from public examinations in that their goal is to inform policy for the education system as a whole rather than to certify and select individuals. They can be used for purposes as varied as allocating scarce resources, monitoring standards, informing teaching, promoting accountability,

7. Eckstein and Noah, op. cit.

8. Vincent Greaney, and Thomas Kellaghan, *Using Examinations to Improve Education: A Study in Fourteen African Countries*, World Bank Technical Paper Number 165, 1992.

9. Idem.

10. Ibid.

11. Eckstein and Noah, op. cit, pp. 15-16.

BEST COPY AVAILABLE

Primary school, Marseilles, France. ©Eric FRANCESCHI/Agence VU

increasing public awareness of the importance of education and informing political debate.

National assessments typically involve the administration of achievement tests to a systematic sample of pupils, but many employ other measurements as well, such as questionnaires for students, teachers and parents aimed at describing the context in which learning takes place. Information on topics such as parental education, teacher preparation and levels of homework assigned can provide valuable guidance in making decisions on national education policy. Continuous monitoring of student performance over time offers a means of identifying and monitoring educational trends.

At least fifty nations carry out national assessments, and the number is growing.[12] Most such

12. Vincent Greaney, and Thomas Kellaghan, *Monitoring the Learning Outcomes of Education Systems*, The World Bank, Washington, DC, 1996, p. 7.

systems are in developed countries, but numerous developing countries, including all but two Latin American nations, have well-established programmes. Chile and Argentina, for example, have comprehensive and well-managed assessment systems, and both are moving towards integrating them into large-scale monitoring and evaluation systems. Chile's current system, established in 1982, tests all primary and secondary students, with results fed back to their respective schools. Argentina's system, initiated in 1994, uses a sampling system.

Community school, Romania. ©Stephane HARTER/Agence VU

International assessments

International assessments are studies that examine samples of students from many different countries and compare results. These studies are organized by international committees that co-ordinate the work of teams of researchers in each country. The organizers devise achievement tests that are then translated into various languages and administered to samples of pupils in the participating countries. Like many national assessments, these international measurement exercises frequently supplement achievement tests with questionnaires for students, teachers and parents in order to gain insights into factors that contribute to learning. Participating countries are typically ranked in numerical order based on the overall performance of their students.

Numerous international assessments have been carried out over the past forty years by the International Association for the Evaluation of Educational Achievement (IEA), headquartered in the Hague. The most extensive and most sophisticated study of this kind was the Third International Mathematics and Science Study (TIMSS) whose results were published in 1997. It sought to measure, compare and explain learning in mathematics and science at three grade levels – primary, middle school and the end of secondary schooling – in forty-one countries and territories. Designers of TIMSS and previous international

BEST COPY AVAILABLE

comparative studies went to great lengths to assure that students in different countries would be taking the same tests and would be judged against international standards of performance.

Because of cost and other factors, most of the participants in past international assessments have been developed countries. When developing nations have participated, the performance of their pupils has tended to lag behind that of their counterparts from other nations. Among Latin American countries, for example, only Colombia and Mexico participated in TIMSS, and only Colombia allowed its scores to be released.

In recent years, however, a number of international assessments have been undertaken with the aim of involving developing countries in cross-national measurements, and this document will examine findings from several of these assessments (see page 25). These include studies by the Southern Africa Consortium for Measuring Educational Quality (SACMEQ), the Monitoring Learning Achievement (MLA) sponsored jointly by UNESCO and UNICEF, the Programme for the Analysis of Educational Systems of the CONFEMEN Countries (PASEC) and the Latin American Laboratory for Assessment of Educational Quality (Laboratorio).

This new generation of international assessments differ from previous ones in some important respects. One of the main breakthroughs of TIMSS was that it permitted each country to be scored both on its own curriculum and on the curricular items common to other countries. Thus TIMSS is both 'international' and 'national' in that it provides a range of test scores that can be used to compare countries, and country specific scores that can be used to assess each country's performance against its own objectives.

The SACMEQ studies adopted the same approach of offering both international and national comparisons and thus set a new standard for large-scale educational survey research in developing countries. The latter approach is consistent with the framework of Jomtien, which urged countries to work toward specific learning objectives for particular grades. Success and failure in meeting these objectives would be judged not in terms of international standards but in relation to standards deemed appropriate to the particular country. The various projects vary in the extent to which they lend themselves to detailed comparison of student achievement across national lines.

Another common feature of the TIMSS and SACMEQ studies is that they provide estimates of national mean test scores along with information that permits judgements to be made about the sampling accuracy of these estimates. These sampling error calculations are essential if one is to make valid claims about the existence of differences between countries and of differences between sub-groups within countries. ■

TYPES OF ASSESSMENTS

Single Country National Assessments (SCNA) – These permit one country to compare its performance with its own objectives. Many countries have conducted such assessments.

Multiple Country National Assessments (MCNA) – These permit each country to compare its performance with its own objectives. Cross-country comparisons are problematic because of different target populations, sampling procedures and testing instruments. Examples include MLA, PASEC and Laboratorio.

Multiple Country International Assessments (MCIA) – These permit countries to be compared on a single test that is designed to fit a common set of curriculum content across countries. Examples include all IEA studies prior to TIMSS.

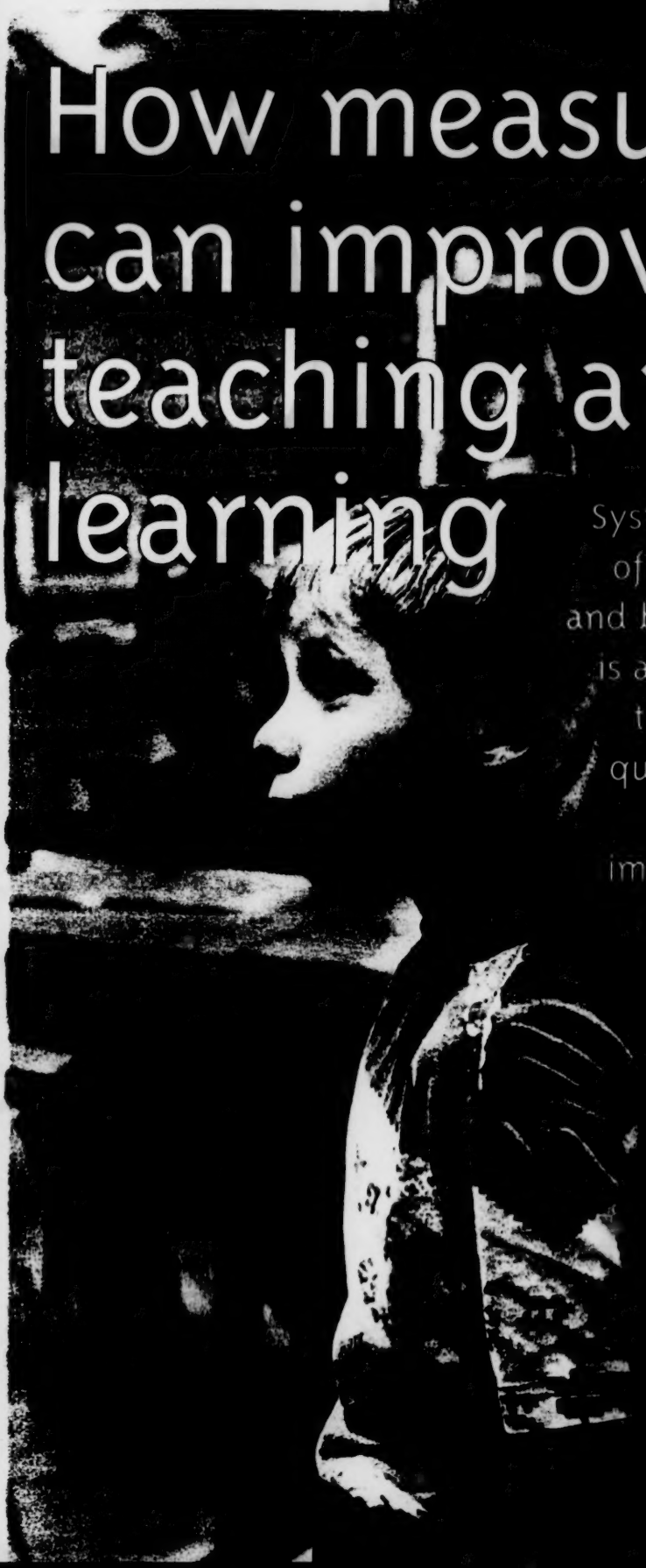
Multiple Country National and International Assessments (MCNIA) – These permit both SCNA and MCIA. In addition, they permit each country to score itself on the curricula of other countries and thus generate a rich analysis of comparative performance. Examples include TIMSS, SACMEQ and PISA.

Rural school, France. ©Eric FRANCESCHI/Agence VU

BEST COPY AVAILABLE

PART II

How measurement can improve teaching and learning



Systematic measurement of student achievement and background variables is a potentially powerful tool for improving the quality of teaching and learning. Such improvement can occur in a number of ways.

Teaching to the test

A considerable body of research has shown that teachers are sensitive to the content of the examinations that their students will be taking. They stress subjects and particular topics that they expect to be on the tests and de-emphasize others. 'Teaching to the test' is intensified when the stakes attached to the examination are high.

For example, in Madagascar subjects such as music and physical education, which do not appear on state examinations, tend not to be taught in primary schools. For the same reason, local languages and practical subjects are undervalued in Zambia, and Ethiopian teachers devote relatively little classroom time to writing, aural skills, the use of reference materials and practical work in science.[13]

While it is regrettable that formal examinations should have the effect of narrowing the range of subject matter that teachers cover in their classrooms, educational policy-makers can use this fact of life to their advantage. Well-constructed examinations that cover a broad range of material that pupils should master can motivate teachers to present appropriate material in their instruction. They can also provide important feedback to teachers on how well students are learning

Alignment of curricula and instruction

Countries design national curricula to embody the knowledge and skills deemed important for pupils to acquire. Even the most finely crafted curriculum is irrelevant, however, if teachers do not use it as the basis for their classroom instruction.

Both public examinations and national assessments can serve the purpose of 'aligning' curriculum and instruction. Working from a coherent set of teaching and learning goals, educational policy-makers can make sure that both the national curriculum and the various measurement devices used to measure student progress reflect these goals. The curriculum then becomes the basis for classroom instruction, while the examinations and assessments serve as a prod for teachers to follow it.[14]

One danger of linking measurement curriculum and instruction is that the assessments can lose their ability to promote reforms. Once in place, curricula can easily become forces of conservatism.

Informing national education policy

Both public examinations and national assessments can provide policy-makers with important feedback on what pupils are learning and not learning. Such feedback can then be used for purposes such as fine tuning the national curriculum or improving the training of new teachers and the in-service training of current ones. Assessment results have inspired officials in Hungary to push reading reform and educators in both Canada and the United States to put more emphasis on science.[15]

As we shall see below, assessment results can be, and frequently are, used to identify achievement problems among particular groups of students. They can also identify country-specific factors – both internal and external to the school – that have positive or negative impact on student achievement. Such information can in turn inform spending priorities and address educational inequities.

Increasing public support for education

Many countries go to great lengths to publicize assessment results as a means of focusing public attention on the importance of education. Favorable results can, of course, be used to enhance public confidence in the state education system and to build support for continued funding, but even unfavorable data can have beneficial effects. Political and educational leaders can point to such data as a way of mobilizing public opinion behind efforts to address educational problems, such as underachievement among particular groups of students.

Limitations of various forms of measurement

While the various types of measurement represent potentially powerful tools for improving teaching and learning, they each have important limitations.

13. Greaney and Kellaghan, 1992, op. cit.

14. Greaney and Kellaghan, 1996, op. cit.

15. Ibid.

322

School-based assessments. A major drawback of school-based assessments is that the standards of evaluation are likely to vary widely from school to school, not only because of teacher bias but because teachers will face pressure from parents and school administrators to inflate students' scores. Such tests can also affect the relationship between teachers and students by increasing the 'judicial' role of the teacher.

Public examinations. Since these tests focus on achievement of individual students, they do not tend to be particularly effective in evaluating learning outcomes for an entire education system, i.e. as a substitute for 'assessment'. In theory, one could simply aggregate individual examination scores in order to make judgements about the system as a whole. In practice, however, aggregating scores poses huge technical problems.

Public examinations are 'norm-referenced', which means that they are designed to show how well pupils do in relation to their peers, not whether they have mastered a particular body of knowledge. Such

VALUE-ADDED ASSESSMENT IN NORTH CAROLINA

The State of North Carolina in the United States has adopted a sophisticated 'value-added' approach to assessment as a means of raising academic standards and enforcing accountability in its public school system.

Under a five-year-old plan known as the 'ABCs of Education', the state's 1.2 million primary-school pupils and secondary-school students are tested each year in reading and mathematics to determine whether they are performing at their grade level. The annual growth in academic performance by each student is then calculated, and each of the state's 2,000 individual schools is then rated on the cumulative growth of its pupils.

Growth goals are set for each school and those that exceed their performance goals by 10 per cent or more receive financial bonuses to be distributed among teachers. Last year more than 80 per cent of teachers in state schools shared bonuses totalling US$120 million. A typical bonus is about US$1,500.

Schools that fall more than 10 per cent below their goals get help from special teaching teams and face the possibility of a state board takeover, as do schools where more than half the pupils are performing below their grade level. The plan also contains provisions for dismissing principals and superintendents who do not help with plans to improve performance.

The most significant feature of the North Carolina plan is the fact that it tracks the growth of individual pupils, even when they move from one school to another. The use of longitudinally tracked data means that the state is not comparing different groups of pupils over time. This feature is especially important for schools in low-income urban areas, where it is not uncommon for the student body to turn over by more than 100 per cent in the course of an academic year.

State officials credit the ABC plan with significantly improving the basic skills of pupils and students. Since its inception the proportion of pupils and students scoring at or above grade level in reading and math has increased from 53 to 69 per cent. Critics, however, say that much of the gain can be attributed to teaching to the test and note that other national examinations have not shown similar gains by North Carolina pupils. Critics also charge that the programme has led to less emphasis on non-tested subjects such as art and physical education.

Controversy has also arisen over issues of equity. Some 52 per cent of African American pupils in Grades 3 to 8 did not score at grade level last year versus only 21 per cent of their white peers. Most of the lowest-performing schools had high proportions of minorities and pupils from low-income families. Critics say that the ABC programme unfairly rewards schools with high proportions of wealthy white pupils. They argue that it should be replaced by a more comprehensive model of accountability using multiple measures of a school's success, not just test scores.

The plan is also at the centre of controversy over recent plans by state officials to end 'social promotion' by retaining pupils who do not perform at grade level on the ABC tests. Educators say that this is an inappropriate use of these tests because they have never been validated for this purpose.

measuring devices tend to concentrate on a relatively narrow band of academic skills required for passage on to the next level of education, and then pay only limited attention to lifeskills and vocational subjects. Thus they have the potential to narrow the curriculum that is actually taught and, by increasing stress on both teachers and students, can have a negative effect on the classroom climate. When high stakes are attached to success on public examinations, the fact that pass rates go up does not necessarily mean that teaching and learning has improved. It could be a consequence of more effective teaching to the test.

The ability of public examinations to inform teaching and learning strategies is also limited by the fact that they are typically taken at the end of an educational cycle – too late to use the information gathered to help those taking the tests.

National assessments. Because these tests are designed to provide information about the larger education system, the individual pupils who take them have no personal stake in doing well. If some pupils do not take the examination seriously, the results may understate the amount of learning that is going on.

National assessments also provide only the most general sort of feedback to schools and classroom teachers regarding strengths and weaknesses of particular students. This is especially true when, as is normal, the assessment involves sampling.

International assessments. Participation in international assessments offers numerous benefits for both developed and developing countries. Such exercises put high-quality measurement devices developed co-operatively at their disposal at relatively low, though by no means trivial, cost. Local officials receive valuable training and hands-on experience in areas such as the development of research questions, sampling techniques, the cleaning of data and how to write up results.

The results of international assessments can help educational authorities to improve their understanding of education systems in general and to draw conclusions about the strengths and weaknesses of their own systems in relation to those of other countries. Such conclusions can be elicited either from scores or sub-scores on international tests that directly reflect the curricula of participating countries or from comparisons of how well a nation's 'intended' curriculum is actually implemented by teachers and mastered by pupils. International comparative studies can also be helpful in identifying the causes of documented differences in pupil performance. Data on factors such as curricula, teacher training, class size, the amount of time spent on school work, parental involvement and socio-economic backgrounds can be helpful for such purposes.

The usefulness of results will be limited, however, to the extent that the particular domains of subjects tested do not coincide with the curricula of particular countries. So-called 'horserace' results showing how students in one country perform in relation to peers in other countries may not be all that useful in designing ways of improving instruction. Assessment results can often be used by political leaders for positive purposes, such as generating additional revenues in order to keep schools competitive with those of other countries. Results can also be used, and in some cases distorted, for self-serving political purposes and lead to consequences that actually hinder educational improvements.

Perhaps the main limitation of measurement devices of all kinds is the failure of political and educational leaders to make the fullest use of the information obtained. 'Examination results are seldom used to provide useful feedback to schools, administrators or curriculum bodies,' write Vincent Greaney and Thomas Kellaghan. 'Thus a good opportunity to effect change is little exploited.'[16] ■

16. Greaney and Kellaghan, 1992, op. cit., p. 2.

Rural school, France. ©B. DESCAMPS/Agence Vu

Community school, Romania
©Stephane HARTER/Agence VU

Primary school, Marseilles, France.
©Eric FRANCESCHI/Agence VU

BEST COPY AVAILABLE

Early childhood education, Cambodia. (Hien Lam Elementary)

BEST COPY AVAILABLE

PART III

Recent research on the quality of education

A new generation of national and cross-national studies has given us important new insights into what pupils in primary schools in developing countries know and are able to do. Some of these studies also provide valuable background information on the context in which learning takes place and discuss factors that contribute to quality teaching and learning. This document will focus on six studies described below.

1. Southern Africa Consortium for Measuring Educational Quality (SACMEQ)

This series of individual country reports analysed the reading levels of pupils in Grade 6 in southern Africa. Five surveys were completed in 1998 (Mauritius, Namibia, Zambia, Zanzibar and Zimbabwe) and two others (Kenya and Malawi) are expected to do so by the end of 1999. A second round of studies involving fifteen countries is scheduled for 2000.

The SACMEQ project was undertaken by Ministries of Education in co-operation with the International Institute for Educational Planning (IIEP) of UNESCO. IIEP provided technical support for the project and facilitated the international workshops at which the research design and data collection methods were developed.

The project's objective is to 'undertake educational policy research with the main aim of generating reliable information that can be used by decision-makers to plan the quality of education.'[17] Another important goal is to enhance the research and evaluation capacity of each of the national education systems.

2. Monitoring Learning Achievement (MLA)

Under this joint UNESCO-UNICEF project, studies of learning achievement are being carried out in forty developing countries in Africa, Asia, the Arab world, the Caribbean, Europe and Latin America. On the basis of the data collected, countries are able to identify factors that promote or hinder learning in primary school, analyse problem areas and develop policy changes and new practices to improve the quality of education.

17. Dhurumbeer, Kulpoo, *The Quality of Education: Some Policy Suggestions Based on a Survey of Schools, Mauritius*. SACMEQ Policy Research: Report No. 1. Ministry of Education and Human Resource Development, Mauritius and the International Institute for Educational Planning, Paris. Foreword.

BEST COPY AVAILABLE

The project is unusual in that, in addition to testing traditional academic subjects in light of each country's national curriculum, it looks at life skills such as knowledge of health, hygiene and nutrition. The MLA studies administer questionnaires to pupils, class teachers, head teachers and parents. As with SACMEQ, a major goal of these studies is to enhance the assessment capacity of the participating countries. It employs a 'critical mass' approach in which members of core groups receive training and then spread out to train other evaluators.

3. Latin American Laboratory for Assessment of Educational Quality (Laboratorio)

The Laboratorio project, with UNESCO sponsorship, is a network of assessment systems in eighteen Latin American countries. Tests are given to representative samples of students in Grades 3 and 4 in each country with the general aim of improving policy development at both the macro and micro levels.

Celebration, Bolivia. © Hien LAM DUC/Agence VU

The first study, carried out in 1997 in thirteen countries, was curriculum-referenced and yielded data showing how achievement related to a wide range of variables, starting with demographics (megacities, urban, rural) and whether sponsorship was public or private. The study also provided new information on the extent to which achievement is a function of expenditures for purposes such as teaching materials or reducing class size, and it looked at various educational practices, such as child-centred v.s. teacher-centred teaching and various assessment strategies. The study evaluated the impact of school policies in areas such as grade retention and principal leadership, nutrition programmes and family and community factors, including parental education and involvement and community support.

Although comparative results are published on the achievement levels of students in each of the countries, differences in sampling and other technical issues mean that the data are most useful in analysing regional and other differences within each participating nation. A second project, currently underway, includes a quasi-longitudinal study, a skill-oriented approach in test development and a test of human development.

4. Third International Mathematics and Science Study (TIMSS)

The most extensive international comparative study to date, TIMSS, sought to measure, compare and explain learning in mathematics and science at three grade levels – primary, middle school and the end of secondary schooling – in forty-one countries and territories. Tests and questionnaires were administered in 1995, and reports at the national and international levels began to be released the following year.

TIMSS is co-ordinated by the International Association for the Evaluation of Educational Achievement (IEA), an international co-operative of research centres and departments of education in more than fifty countries with headquarters in the Netherlands.

Tests are designed by task forces from participating countries, but each nation is responsible for conducting its own data collection and analysis in co-operation. The process is overseen by an international steering committee that provides technical assistance, monitors the sampling process and oversees the analysis and reporting of findings.

IEA has sponsored more than fifteen studies of cross-national achievement in curricular areas such as mathematics, science, language, civics and reading. The IEA conducted its First International Mathematics Study (FIMS) in 1964 and the Second International Mathematics Study (SIMS) in 1980-82. The first and Second International Science Studies

High School, Connecticut, USA. © Alain BIZOS/Agence VU

TIMSS is a particularly ambitious project not only because of the large number of countries involved but because of the scope of the research. In addition to administering tests in the two subjects at the three grade levels, researchers asked students, teachers and school administrators to fill out questionnaires aimed at producing a better understanding of how contextual factors affect achievement. For example, students were asked about their attitudes towards the schooling, and teachers about how they structured their classrooms. In some cases, countries supplemented these core activities with curriculum analysis, videotaped observations of classrooms, case studies and assessments of sub-sets of students given hands-on mathematics and science tasks.

TIMSS represents the continuation of a long series of studies conducted by the International Association for the Evaluation of Educational Achievement (IEA). Since its inception in 1959, the (FISS and SISS) were carried out in 1970-71 and 1983-84, respectively. Since the subjects of mathematics and science are related in many respects and since there is broad interest in many countries in students' ability in both mathematics and science, the third studies were conducted together as an integrated effort.

5. Programme for the Analysis of Educational Systems of the CONFENEM Countries (PASEC)

Established in 1991 by the ministers of education of French-speaking countries in sub-Saharan Africa, PASEC has worked to use assessment as a way of identifying efficient models of schools. Studies have been carried out in nine countries: Burkina Faso, Cameroon, Central African Republic, Congo, Djibouti, Côte d'Ivoire, Madagascar, Mali and Senegal.

EVALUATION OF A NON-FORMAL EDUCATION PROGRAMME IN GHANA

In 1996, the Bimoba Literacy and Farmers Co-operative Union (BILFACU) a small organization in Ghana working in the areas of literacy, food security, women and micro-credit, decided to work with the non-governmental organization ActionAid and its REFLECT literacy programme.

Situated in the isolated East Mamprusi District where the main town has no electricity, and the road no tarmac, the area is inaccessible in the rainy season. The border with Togo is near and people share the common language of Moor. Few non-governmental organizations work in the area where cotton is the main industry and where food insecurity and environmental degradation are rampant.

The REFLECT programme aims at bringing about literacy, a change of attitude and development in communities at low cost. BILFACU started by selecting facilitators in consultation with the community. Criteria were that they be literate in Moor and English, that they have a spirit of volunteerism and that they be respected members of their community. All facilitators were trained teachers and one was also a village chief.

After a nine-day training course in November 1997 and a refresher in February 1998 when courses began, groups started meeting three times a week in the afternoons. In the rainy season they shifted to evenings and reduced meetings to once a week. Some 127 learners were enrolled, mostly women. Learning materials, supplied by BILFACU, included exercise books, silk screen printing materials, plywood blackboards, chalk, pencils, wooden box for materials.

The evaluation took place from 8-10 March 1999. The objective of this self-evaluation was to help everyone, including participants, facilitators, BILFACU and ActionAid, to see what was achieved and how to proceed.

Evaluators reported the following results :

REFLECT Villages	Initial enrolment	Current learners	Retention rate
JIIRIK	30	10	33%
NAKORUK	30	27	90%
NAAUK	28	22	79%
NABAUK	18	17	94%
TOJING	21	13	62%
TOTALS	**127**	**79**	**63%**

Most drop-outs occurred within the first few weeks of the programme, and the retention in four of the villages compare favourably with most literacy programmes which often experience a one-third drop-out rate. It was unclear why Jiirik had such a low retention rate compared to the other circles, particularly as some records seemed to be very good. Enrolment had been a one-off exercise and new recruits were not encouraged. The evaluation team found that new learners were waiting to start with new study circles in two out of the five villages visited.

Attendance had followed a seasonal pattern (depending on agricultural work and funerals, etc.). It was considered very realistic to reduce the number of meetings to once a week in these periods, as facilitators could become demotivated if they felt they had lost time from their own farms for a very small number of learners. Irregular attendance due to personal responsibilities and problems was also reported, and again was felt to be typical. The effects of this irregular attendance was seen on the poor literacy results (see below).

The facilitators' attendance was excellent. They did not receive allowances for attending training and no incentives apart from a bicycle, rain coat, wellington boots and solar lamp (all as working tools). The evaluation team felt that this was the result of regular monitoring visits, and frequent and intensive training as well as the community dynamics. Their high educational level and the fact that most were teachers receiving a regular salary may also have meant they found it easier. In general, facilitators displayed a high degree of ownership of the programme and were initiators rather than just passive recipients of training.

The rate of learning, however, was disappointing. Evaluators found that, generally, reading, writing and numeracy skills of participants were poor. Causes were identified as ranging from lack of time, lack of books and poor lighting; poor training of facilitators in this area; lack of interest in Moor and the desire of participants to learn English, especially spoken English.

The evaluators, however, found empowering effects of the REFLECT programme on learners. For instance, women had established the right to sit outside at night with men — and not be "sacked" by their husbands, and dared attend the medical clinic without a male escort. Being able to find their own cards, they were more confident in dealing with clinic staff. A woman participant was elected as Chair of Literacy Committee, the first woman to lead both women and men within living memory. Parents started an informal school and built two blocks themselves. They now have Primary One and Primary Two classes.

Participants started moulding mud bricks to rehabilitate an old literacy centre to continue with the circle meetings. Stone walls were built to prevent soil erosion, and people decided to try and store food for the "hungry season", as opposed to the traditionally extravagant post-harvest feasting. A tree reservation had been started, and was observed to be very green and bushy a year later, and money was collected to regrade the road.

BEST COPY AVAILABLE

Each of these countries has carried out at least two major assessment projects focusing on achievement in mathematics and French in Grades 2 and 5, with tests administered at the beginning and the end of the academic year. The PASEC research is notable in that it has explored the impact on achievement of sixteen in-school factors (such as teacher training, class size and availability of textbooks) and eight environmental influences (including parental education, distance to school and language spoken at home). Researchers expect that their analysis of the causal effect of these different variables will provide policy-makers with important insights into how to improve efficiency and effectiveness of educational systems.

The assessment activities have been carried out by national teams in each country with support from researchers from northern countries. A major goal of the project has been to build up the capacity of each nation to carry out on-going assessments.

In five of the countries the same tests were used, thus making results comparable.

6. Programme for International Student Assessment (PISA)

A new, regular survey of 15-year-olds was launched in 1998 by the twenty-nine member countries of the Organisation for Economic Co-operation and Development (OECD) and Brazil, China, Latvia and the Russian Federation. The programme aims to assess whether education systems are providing the tools for continued learning that today's young people will need over the course of a lifetime for full participation in society.

PISA uses international comparative surveys of student achievement and best-practice analysis to produce policy-oriented and internationally comparable indicators of student achievement on a regular and timely basis.

Bringing together scientific expertise on the basis of shared, policy-driven interests, participating countries want to produce a method of assessing students that is valid across countries, that is strong at measuring relevant skills and that is based on authentic life situations. The first PISA assessment will take place in 2000. Thereafter assessments will occur every three years. Three 'domains' reading literacy, mathematical literacy and scientific literacy form the core of each cycle, but two-thirds of testing time in each cycle will be devoted to a 'major' domain, assessed in depth. Major domains are reading literacy in 2000, mathematical literacy in 2003 and scientific literacy in 2006.

Samples of between 4,000 and 10,000 students will be assessed in each country. A sample of students in each country will complete a variety of pencil and paper tasks. They will also complete a questionnaire about their background and attitudes.

PISA will not only assess the knowledge and skills of students but also ask them to report on their own, self-regulated learning, their motivational preferences and their preferences for different types of learning situations. ■

PART IV

Broad findings of recent research

Rural school, Mali. ©Stanley GREENE/Agence VU

BEST COPY AVAILABLE

As with similar research in the past, the six studies listed above offer a number of broad insights into the status of educational achievement by primary-school pupils in countries throughout the world. These include the following:

The low level of achievement in developing countries

Although there are not a lot of cross-country data on how well pupils in developing countries perform in relation to their counterparts in industrialized nations, the evidence that does exist suggests that the level of performance of these pupils is distressingly low, both by global and national standards.

International standards. Results from the most recent major international comparative study, the Third International Mathematics and Science Study (TIMSS), show that the students in Grade 7 in the few developing countries that participated had the lowest scores in both the mathematics and science sections (Figures 3 and 4). Data also show that the proportion of girls in Grades 7 and 8 who attained what was judged to be a satisfactory score in the mathematics test ranged from 22 per cent for South Africa and 25 per cent in Colombia to 73 per cent in Singapore (Table 1).

After surveying recent efforts to assess learning levels in South Asia, Jim Irvine, Regional Education Adviser for UNICEF in Bangkok, concluded that 'the general picture which has emerged in each country is disappointing.' In general, he said, 'these studies have shown that a substantial proportion of 11-12 year-old children drawn from nationally representative samples could not demonstrate mastery of some defined basic literacy, numeracy and life skills competencies, even though many of the children had been in school for five years or more.'

BEST COPY AVAILABLE

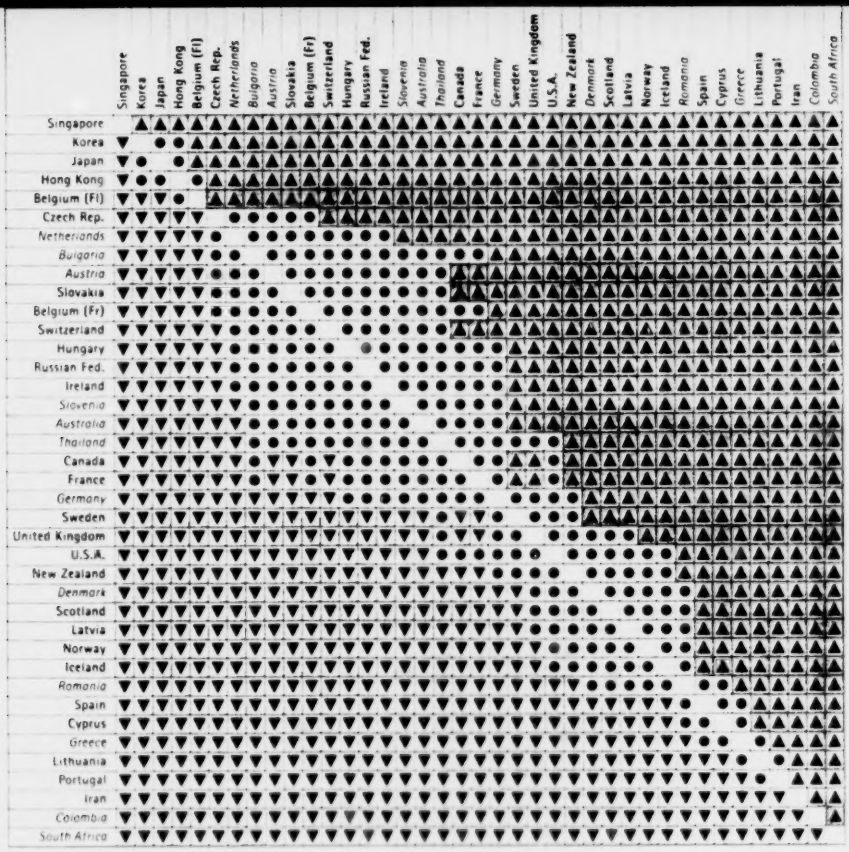


Figure 3

Ranked TIMSS participating countries according to the mean score in mathematics test - Grade 7

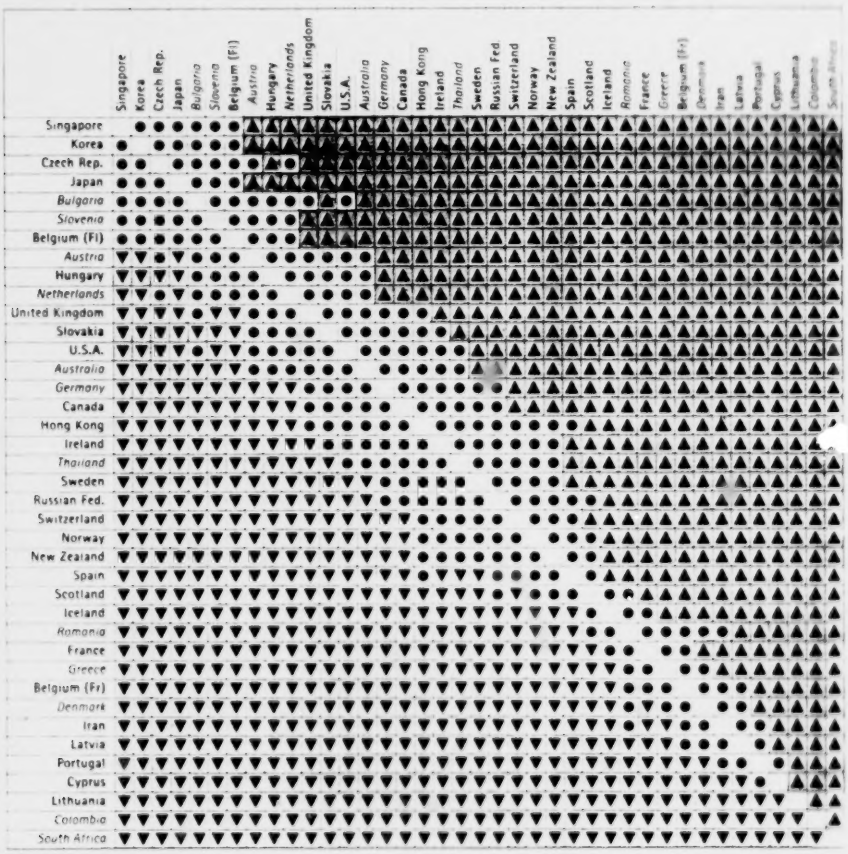


Figure 4

Ranked TIMSS participating countries according to mean score in science test - Grade 7

Note: Countries are ranked according to their mean score in the tests. Each row indicates the result of inter-country comparison. For example, the math's mean achievement for the Flemish community of Belgium is significantly lower than that for Singapore, not significantly different from that for Hong Kong, and significantly higher than that of Norway. Countries shown in italics did not satisfy one or more guidelines for sample participation rates, age/grade specifications, or classroom sampling procedures.

▲ The mean score is significantly higher in the country in row compared to the country in column.

● The mean score is significantly higher in the country in column compared to the country in row.

▼ The mean scores are not significantly different.

Source: International Association for the Evaluation of Educational Achievement (IEA)/ TIMSS, 1994-95.

BEST COPY AVAILABLE

ERRATUM

Page 32

The caption should read :

▲ : The mean score is significantly higher in the country in row compared to the country in column.

▼ : The mean score is significantly lower in the country in row compared to the country in column.

● : The mean scores are not significantly different.

Table 1

Percentage of items correct on TIMSS by grade, by subject and by gender

Country/Territory	Grade 7								Grade 8							
	Mathematics				Sciences				Mathematics				Sciences			
	Boys		Girls		Boys		Girls		Boys		Girls		Boys		Girls	
Australia	52	(1.2)	53	(1.0)	54	(1.0)	54	(0.7)	57	(1.2)	59	(1.1)	61	(1.0)	59	(0.8)
Austria	55	(1.1)	56	(0.8)	56	(0.9)	55	(0.7)	63	(0.8)	61	(1.2)	63	(0.8)	60	(0.8)
Belgium (Fl)	65	(1.1)	66	(1.1)	59	*(0.7)	55	(0.7)	65	(2.0)	66	(1.9)	62	(1.7)	59	(1.5)
Belgium (Fr)	56	(1.0)	53	(1.1)	47	*(0.8)	43	(0.7)	59	(1.1)	58	(1.0)	52	(1.0)	49	(0.7)
Bulgaria	54	(1.5)	56	(1.9)	56	(0.9)	57	(1.2)	59	(1.4)	61	(1.2)	61	(1.1)	62	(1.1)
Canada	52	(0.6)	52	(0.6)	55	(0.6)	53	(0.5)	59	(0.7)	59	(0.6)	60	(0.6)	58	(0.6)
Colombia	27	(0.8)	25	(1.0)	37	*(0.9)	33	(0.8)	30	(1.6)	29	(0.9)	40	(1.4)	37	(0.8)
Cyprus	42	(0.6)	42	(0.5)	40	(0.6)	40	(0.5)	47	(0.6)	48	(0.6)	46	(0.4)	47	(0.6)
Czech Rep.	58	(1.1)	57	(1.3)	60	*(0.7)	56	(0.9)	67	(1.0)	64	(1.3)	67	*(0.8)	61	(1.1)
Denmark	45	(0.7)	43	(0.7)	46	*(0.6)	42	(0.6)	54	*(0.8)	50	(0.9)	54	*(0.6)	48	(0.8)
France	52	(0.9)	50	(0.8)	48	*(0.7)	44	(0.7)	62	(0.8)	61	(0.9)	55	*(0.7)	52	(0.7)
Germany	49	(1.3)	49	(1.1)	55	(1.0)	51	(0.9)	54	(1.3)	54	(1.2)	59	(1.2)	57	(1.0)
Greece	40	(0.7)	41	(0.6)	45	(0.7)	44	(0.5)	51	(0.9)	48	(0.7)	54	*(0.6)	50	(0.6)
Hong Kong	66	(2.2)	64	(2.0)	54	(1.5)	52	(1.2)	72	(1.7)	68	(1.7)	60	*(1.1)	55	(1.1)
Hungary	53	(0.9)	54	(1.0)	57	(0.8)	54	(0.7)	61	(0.8)	62	(0.8)	63	*(0.7)	59	(0.7)
Iceland	43	(0.7)	43	(0.7)	47	(0.9)	45	(0.6)	49	(1.3)	50	(1.3)	53	(1.2)	51	(0.9)
Islamic Rep. of Iran	33	(0.7)	31	(0.7)	43	(0.7)	40	(0.9)	39	(0.8)	36	(0.8)	49	*(0.8)	45	(0.8)
Ireland	55	(1.5)	52	(1.1)	54	*(1.0)	50	(0.8)	60	(1.6)	58	(1.4)	60	(1.3)	57	(1.0)
Israel	-	-	-	-	-	-	-	-	61	(1.5)	55	(1.5)	61	*(1.2)	54	(1.1)
Japan	68	(0.6)	66	(0.4)	60	*(0.4)	58	(0.3)	74	(0.5)	73	(0.4)	67	*(0.5)	64	(0.4)
Rep. of Korea	68	(0.8)	65	(0.9)	63	*(0.5)	59	(0.6)	73	*(0.6)	70	(0.7)	67	*(0.5)	64	(0.5)
Kuwait	-	-	-	-	-	-	-	-	29	(1.1)	31	(0.5)	39	(1.1)	47	(0.7)
Latvia	44	(1.0)	44	(0.8)	43	(0.7)	40	(0.6)	52	(1.0)	51	(0.8)	52	*(0.8)	48	(0.6)
Lithuania	37	(0.9)	39	(0.9)	38	(0.7)	37	(0.8)	48	(1.1)	49	(1.0)	51	*(0.8)	47	(0.8)
Netherland	56	(1.3)	55	(1.1)	57	(0.9)	55	(0.8)	61	(1.8)	59	(1.6)	64	(1.2)	60	(1.1)
New Zealand	46	(1.0)	46	(0.9)	51	(0.8)	49	(0.7)	55	(1.4)	53	(1.3)	60	(1.0)	56	(1.0)
Norway	45	(0.8)	43	(0.8)	51	(0.7)	49	(0.8)	54	(0.6)	53	(0.6)	59	(0.6)	56	(0.4)
Portugal	37	(0.7)	36	(0.6)	43	*(0.5)	39	(0.5)	44	(0.8)	42	(0.7)	52	*(0.7)	48	(0.6)
Romania	43	(0.9)	43	(0.9)	46	(0.8)	44	(0.8)	49	(1.1)	49	(1.0)	51	(0.9)	49	(0.9)
Russian Fed.	53	(1.2)	53	(0.8)	52	(1.0)	48	(0.7)	59	(1.4)	61	(1.3)	60	(0.9)	57	(0.7)
Scotland	45	(1.1)	44	(0.9)	50	(0.9)	47	(0.8)	53	(1.7)	50	(1.3)	57	(1.2)	53	(0.9)
Singapore	73	(1.4)	73	(1.6)	62	(1.4)	61	(1.5)	79	(1.1)	79	(1.0)	71	(1.2)	69	(1.1)
Slovakia	55	(1.1)	54	(0.8)	57	*(0.8)	52	(0.6)	63	(0.9)	62	(0.8)	62	*(0.6)	57	(0.7)
Slovenia	53	(0.8)	52	(0.8)	59	(0.6)	56	(0.6)	62	(0.8)	60	(0.7)	64	*(0.6)	59	(0.7)
South Africa	24	(1.4)	22	(0.8)	27	(1.3)	25	(0.9)	25	(1.7)	22	(1.0)	28	(1.8)	25	(1.2)
Spain	43	(0.6)	42	(0.7)	51	*(0.6)	47	(0.5)	52	(0.7)	50	(0.7)	58	*(0.5)	54	(0.5)
Sweden	47	(0.7)	47	(0.8)	52	(0.6)	50	(0.7)	56	(0.8)	56	(0.8)	60	*(0.6)	57	(0.6)
Switzerland	54	(0.6)	52	(0.6)	52	*(0.5)	48	(0.5)	63	(0.8)	61	(0.7)	58	*(0.6)	54	(0.5)
Thailand	51	(1.2)	52	(1.4)	53	(0.8)	52	(0.9)	56	(1.4)	58	(1.7)	57	(0.9)	58	(1.0)
United States	48	(1.3)	48	(1.3)	55	(1.3)	53	(1.1)	53	(1.2)	53	(1.1)	59	(1.0)	57	(1.0)
United Kingdom	49	(1.4)	45	(1.0)	57	(1.0)	54	(0.9)	53	(1.3)	53	(0.9)	63	(1.0)	60	(0.7)

* The difference is significant at 0.05 level.

Note: Countries shown in italics did not satisfy one or more guidelines for sample participation rates, age/grade specifications, or classroom sampling procedures. Figures in parenthesis indicate the standard errors.

Source: International Association for the Evaluation of Educational Achievement (IEA)/ TIMSS, 1994-95.

Likewise, after examining how Latin American countries fared in recent international assessments, Laurence Wolff concluded that these countries 'have consistently scored well below those of North America, Europe and Asia.'[18]

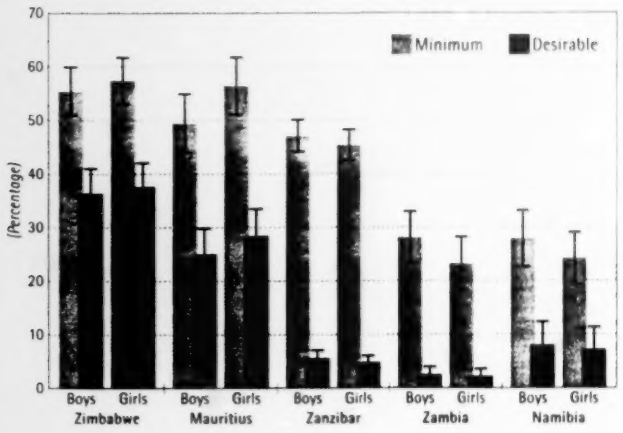
National Standards. Perhaps more importantly, pupils in developing countries are frequently failing to meet the performance standards of their own ministries of education.

The SACMEQ reading literacy study conducted in five countries of southern Africa found that pupils were generally performing poorly when judged by the standards of mastery set down by the respective ministries' own reading experts and selected Grade 6 teachers. Figure 5 shows the proportion of pupils reaching 'minimal' and 'desirable' levels of mastery. In only two of the five studies, Mauritius and Zimbabwe, did at least half of their pupils attain the minimum level of reading fluency, and only in Zimbabwe did at least a third of pupils attain the 'desirable' level.

18. Laurence, Wolff, *Educational Assessments in Latin America: Current Progress and Future Challenges*, Occasional Paper Series No 11, Partnership for Educational Revitalization in the Americas, Washington DC.

Figure 5

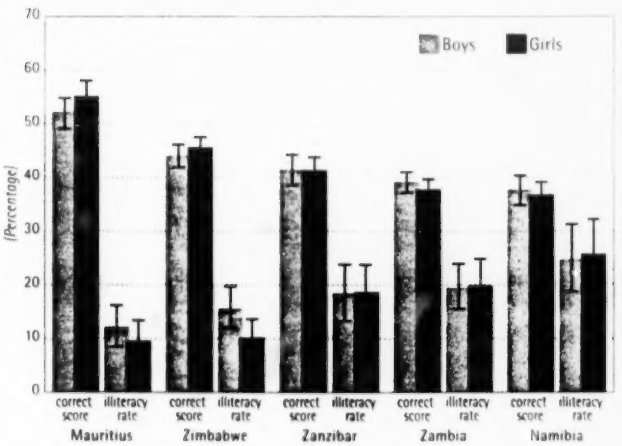
Percentage of pupils reaching minimum and desirable mastery levels in SAQMEC reading-literacy test by gender



Source: SACMEQ Policy Research: Reports No 1-5. IIEP. UNESCO. 1998.

Figure 6

Percentage of items correct on reading literacy test for grade 6 pupils and "illiteracy rate" by gender in 5 SACMEQ studies



Source: SACMEQ Policy Research: Reports No 1-5. IIEP. UNESCO. 1998

Note: the small bars in the charts indicate the sampling errors that allow to construct confidence intervals around the means scores

Figure 6 shows that, with the exception of Mauritius, where Grade 6 pupils correctly answered 53 per cent of the test items, none of the countries reached the level of 50 per cent correct. Researchers also defined an 'illiteracy rate' as the proportion of pupils scoring lower than the score most likely to be obtained by a pupil who guessed on all fifty-nine items. In the cases of Namibia and Zambia, 20 per cent of all pupils qualified as illiterate.

In Nigeria when researchers administered cognitive tests of literacy numeracy and life skills to fourth graders as part of the Monitoring Learning Achievement project, the results were described as 'generally poor'. The mean percentage of scores was 32 per cent in numeracy, 25 per cent in literacy and 33 per cent in life skills. In one test item pupils were instructed to copy a five-line passage into a given space. Only 8 per cent of them were able to do so accurately, and 40 per cent were unable to copy a single word or punctuation mark.[19]

Change in pupil performance since Jomtien

Very few longitudinal data are available to indicate whether the academic performance of primary-school pupils in developing countries has improved since the Education for All Conference in Jomtien in 1990.

Evidence from some countries suggests that it has. In India, for example, the District Primary Education Programme, launched in 1994 in forty-two districts across seven states, established benchmarks for achievement in language and math at the conclusion and the penultimate year of the primary schooling cycle. An assessment was then carried out using a sample of 66,831 pupils to determine whether their achievement had improved. The results showed that substantial majorities of districts improved performance in both subject areas and at both grade levels.[20]

Some countries have reported no change. As part of its SACMEQ project, Zimbabwe set out to determine whether there had been any recent improvement in the literacy levels of Grade 6 pupils. Researchers identified thirty-six test items that were common to a 1991 IIEP study and the 1995 SACMEQ survey, and used these items to compare the performance of pupils. The results are shown in Figure 7. The researchers reported 'no significant changes in the mean scores of Grade 6 pupils.' Researchers said that, since the base scores were 'already very low', the lack of progress warranted a 'major enquiry' by the Ministry of Education and Culture.[21] Other scholars have suggested, however, that, given structural changes that were affecting the country during that period, the fact that scores did not decline should be viewed as a positive sign.

19. Kiem M., Chiejine, *MLA Project Assesses Young Students in Nigeria*, UNICEF Education Update, January 1999, Volume 2, Issue 1.

20. DPEP Core Resource Group, National Council of Educational Research and Training, *Mid-Term Assessment Survey: An Appraisal of Students' Achievement*, New Delhi, October 1998.

21. Thomas, Machingaidze, Patrick Pfukani and Sibangani Shumba, *The Quality of Education: Some Policy Suggestions Based on a Survey of Schools, Zimbabwe*, International Institute for Educational Planning, Ministry of Education and Culture, Zimbabwe.

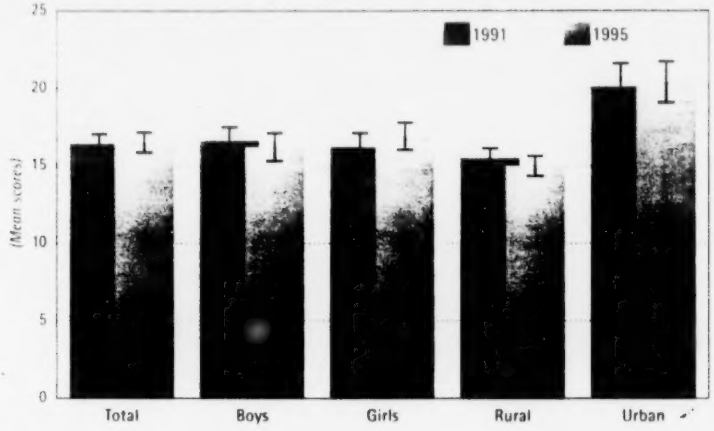


Figure 7

Mean performance of Zimbabwe's Grade 6 pupils on the 36 items that were common to the reading tests administered in 1991 and 95

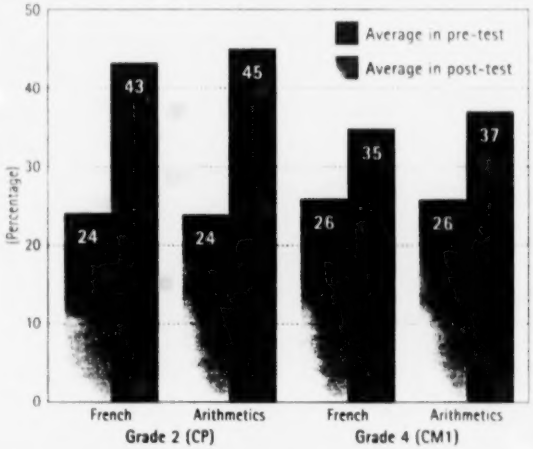
Source: SACMEQ Policy Research: Reports No. 3. IIEP. UNESCO. 1998.

The Zimbabwe study looked at different groups of Grade 6 pupils in the two years. A better estimation of change in achievement can be obtained when assessments are administered to the same groups of students at different points in time. The PASEC project undertook such an investigation by testing Senegalese students in Grades 2 and 4 in November 1995 and May 1996. As seen in Figure 8, pupils in both grades gained in learning achievement during the school year, but the gains were substantially higher for Grade 2.

Figure 8

Learning gain of the same Grade 2 and Grade 4 Senegalese pupils in French and arithmetic

	Grade 2 (CP)		Grade 4 (CM1)	
	French	Arithmetics	French	Arithmetics
Pre-test				
Average	24.1	23.9	25.9	25.8
Standard deviation	20.3	18.9	16.2	16.3
Post-test				
Average	43.2	45.0	34.8	37.0
Standard deviation	25.2	23.2	16.9	16.6
Absolute growth	19.1	21.1	8.9	11.2
relative gross growth	79%	88%	34%	43%
relative net growth	25%	28%	12%	15%



Source: Barrier, E. et al. (1997). Evaluation du système educatif sénegalais. Enseignement élementaire. PASEC-INEADE-CIEP. 1997.

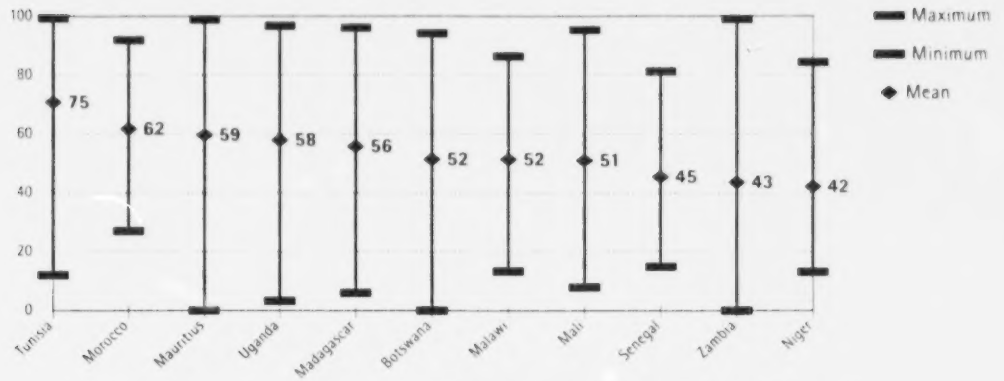
Differences within countries exceed those between them

One important lesson from cross-national studies is that achievement differences between regions *within* countries tend to surpass those *between* countries. To be sure, large-scale projects such as TIMSS report huge differences in the median scores of countries at the top and those at the bottom and, as noted above, between developed and developing nations. Within regions of the world, however, and among nations that have common cultural, linguistic, socio-economic and other features, differences between nations are small. The following figure, drawn from the most recent MLA studies, shows that the variations in performance between pupils *within* countries are larger than those *between* countries.

BEST COPY AVAILABLE

Figure 9

Range of pupils' performance in 11 MLA countries



Source: V. Chinapah, et al. *With Africa for Africa. Towards Quality Education for All. Draft Regional Report. EFA 2000 Assessment. MLA Project*, 1999.

In their study of schools in thirteen Latin American and Caribbean countries, researchers from the Laboratorio concluded that, with the important exception of Cuba, achievement differences between the nations 'are not highly significant'.[22] In Figures 10 and 11 the achievements of each country are shown as standardized scores with a mean of 250 points. The scores are distributed around the mean with a standard deviation of 50. The horizontal bars cover achievements of the middle half of the students (between 25 and 75 per cent), with the mid-point corresponding to the dividing line between the top and bottom 50 per cent of scores. Also shown is the indication of the first and ninth decile, which show the extreme scores in each country. The graph shows that, Cuba excepted, there are no significant differences in terms of the mean and range of test achievement scores among the Latin American countries with regard to Grade 3 language even though the distribution of the results shows a relatively high heterogeneity within the various countries. Similar results are observed for Grade 4 language and mathematics.

22. *Latin American Laboratory for the Assessment of Quality in Education: First International Comparative Study*, UNESCO Santiago, Chile, 1998.

Figure 10

Distribution of third-grade language scores in Latin American countries

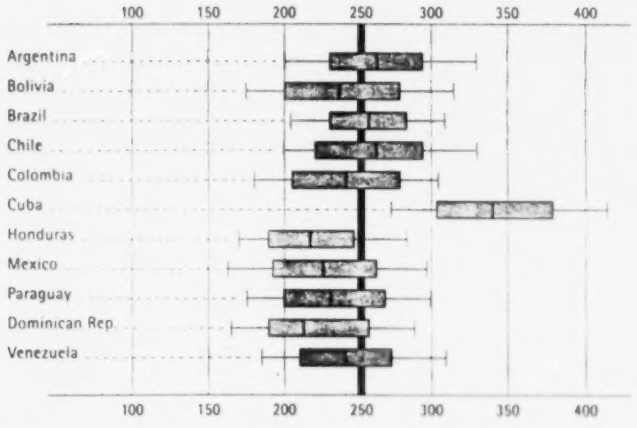
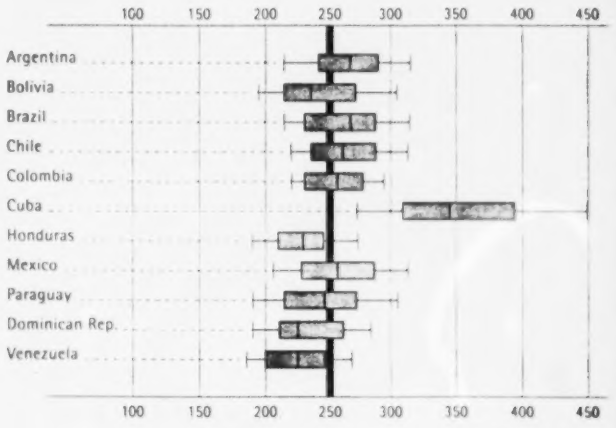
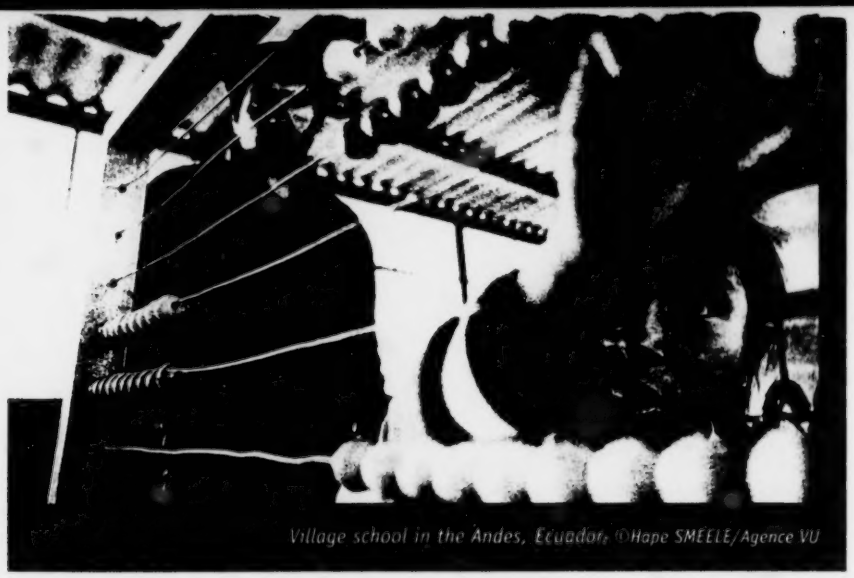


Figure 11

Distribution of fourth-grade mathematics scores in Latin American countries



Source: Latin American Laboratory for the Assessment of Quality in Education. First International Comparative Study of Language, Mathematics, and Associated Factors in Third and Fourth Grades. UNESCO, 1998.



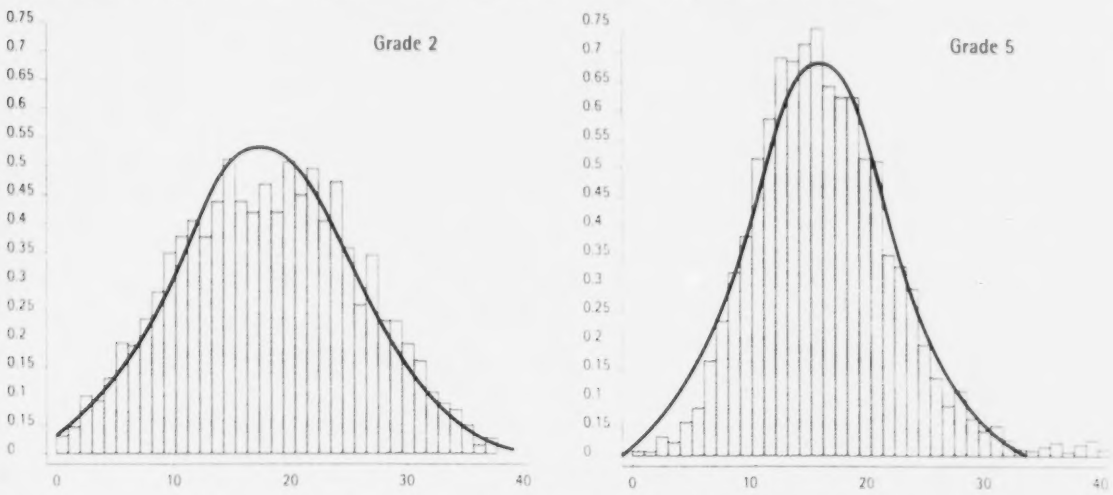
Village school in the Andes, Ecuador. ©Hape SMEELE/Agence VU

At the same time, significant achievement variations can usually be found among the different regions, types of schools and groups of students within a particular country. Understanding these variations is crucial if assessments are to be used as a basis for policies aimed at improving pupil achievement. The purpose of monitoring learning achievement is not only to promote a high level of overall performance but to eliminate, or at least to minimize, achievement disparities between various groups of schools.

Large differences in achievement scores for students attending the same level of education suggest that school systems are not addressing learning obstacles related to the backgrounds and particular needs of students. Figure 12 gives the distribution of scores in mathematics for Grade 2 and Grade 5 pupils in Cote d'Ivoire and shows how the heterogeneity in pupils achievements tends to be reduced as they progress through the system. The flatter graph on the left shows a high dispersion, i.e. high heterogeneity, in pupil achievement in Grade 2, while a relatively high concentration of pupils scored around the average in Grade 5. This pattern can be observed in other PASEC countries studied, including Burkina Faso, Cameroon and Senegal. While one possible interpretation of this trend is that some educational disparities are narrowing as pupils progress at school, another possibility is that the school selection process has excluded low achievers and retained a relatively small number of better students.

Figure 12

Distribution of scores in mathematics for Grade 2 and Grade 5 pupils in Côte d'Ivoire



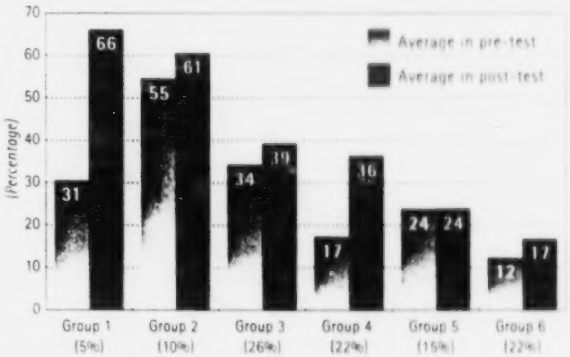
Source: Programme d'Analyse des Systemes Educatifs (PASEC). Les facteurs de l'efficacité de l'enseignement primaire: données et résultats sur cinq pays d'Afrique et de l'Ocean Indien. CD-ROM (Version de mars 1999)

Figure 13 shows the disparities that PASEC researchers found in the performance of Senegalese pupils who were learning French. While variation in achievement scores among pupils is to be expected in most teaching situations, it becomes a problem if a relatively large proportion of pupils falls into the category of under-achievers. Such pupils may be at risk of being unable to pursue lifelong learning or to integrate effectively into society and the world of work.

Figure 13

Percentage of correct items and learning gain in French language for Senegalese Grade 4 (CM1) pupils

	Group 1 (5%)	Group 2 (10%)	Group 3 (26%)	Group 4 (22%)	Group 5 (15%)	Group 6 (22%)
	From a good level, strong progress	From a very good level, fair progress	From a good level, weak progress	From a low level, strong progress	From a fair level, no progress	From a lowest level, low progress
Pre-test						
Average in pre-test	30.5	54.5	34.2	17.2	23.7	12.2
Standard deviation	10.1	13.3	9.7	10.6	9.2	8.9
Post-test						
Average	66.1	60.5	39.2	36.3	23.8	16.6
Standard deviation	9.4	10.9	9.4	8.7	9.0	7.2
Absolute growth	35.6	6.0	5.0	19.1	0.1	4.4
Relative gross growth	117%	11%	15%	111%	0%	36%
Relative net growth	51%	13%	8%	23%	0%	5%



Source Barrier, E. et al. (1997). Evaluation du systeme educatif senegalais. Enseignement elementaire. PASEC-INEADE-CIEP. 1997.

The MLA project in Slovakia also found important regional differences. Pupils in Bratislava showed statistically higher proficiency in Slovak language and life skills than those in both East Slovakia and West Slovakia.[23] Similarly, a study in Sri Lanka found that pupils were performing 'at a very low level' in the districts of Moneragala, Kegalle and Ratnapura. The authors called for giving priority to these districts in the allocation of facilities, supplies and teacher training.[24]

Regional disparities in pupils performance

In many countries there are wide disparities in academic achievement between pupils in various geographic regions. Such differences result from numerous factors, including characteristics of the population, and the human and financial resources available to support schools.

In their studies in southern Africa, SACMEQ researchers found major differences in reading between urban and rural areas except in Zambia. This result can be seen in Figure 14. Figure 15 shows the range of disparities in achievement between the highest and lowest scoring regions in five SACMEQ countries. Notable are the wide regional variability in Namibia and the apparent homogeneity in Zambia.

When they looked at reading levels among Grade 6 pupils, SACMEQ analysts in Namibia found that 'at the overall national level the picture was rather gloomy,' with only 26 per cent of learners reaching 'minimal' level of mastery. Closer scrutiny of the data showed that the proportion of pupils attaining minimal mastery ranged from 5 per cent in the region of Katima Mulilo to 60 per cent in Windhoek. Similar distributions were found among the 8 per cent of students who scored at a 'desirable' level, with some regions producing virtually no such students.[25]

The SACMEQ study in Mauritius found that, by and large, the state has succeeded in providing 'reasonably adequate' material inputs to schools and 'ensuring that there has been an equitable distribution of these inputs among districts'. By contrast, researchers found 'considerable inequity' across districts when it came to human resources, notably the quality of teachers and school heads.[26]

23. Stefan Matula, Stanislav Fila, Rudo Hurban, Eva Smikova, Ludmila Matulova. *Assessment of Learning Achievement of Grade 1 Students in Slovakia: A Preliminary Report*. Defence for Children International – Slovak Stion, Bratislava, May 1995.

24. Wijesuriya, George. *Participation and Performance in Primary Education*. National Institute of Education, Maharagama, Sri Lanka.

25. Friedheim Voigts. *The Quality of Education: Some Policy Suggestions Based on a Survey of Schools, Namibia*. International Institute for Educational Planning and Ministry of Basic Education and Culture, Namibia. p. 64.

26. Kulpoo, op. cit., pp. 61-3.

Figure 14

Difference of mean score between urban and rural area in five SACMEQ countries

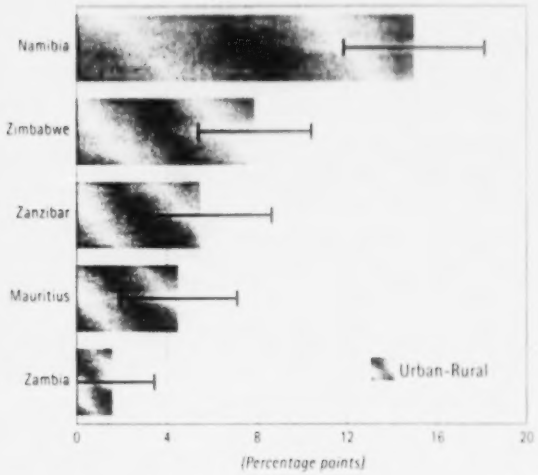
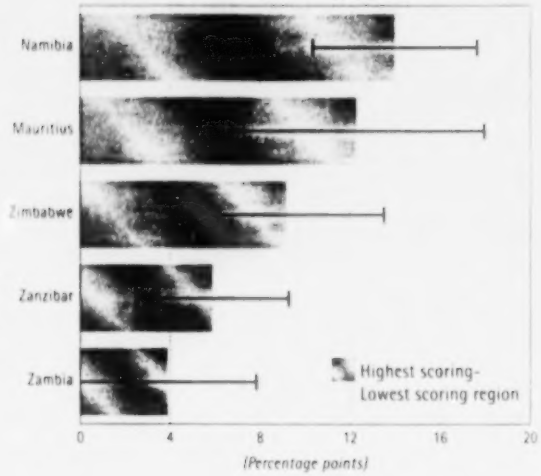


Figure 15

Difference of mean score between the highest-scoring and the lowest-scoring regions in five SACMEQ countries



Source: SACMEQ Policy Research: Reports No. 1-5. IIEP, UNESCO. 1998.

Achievement differences between types of schools and groups of pupils

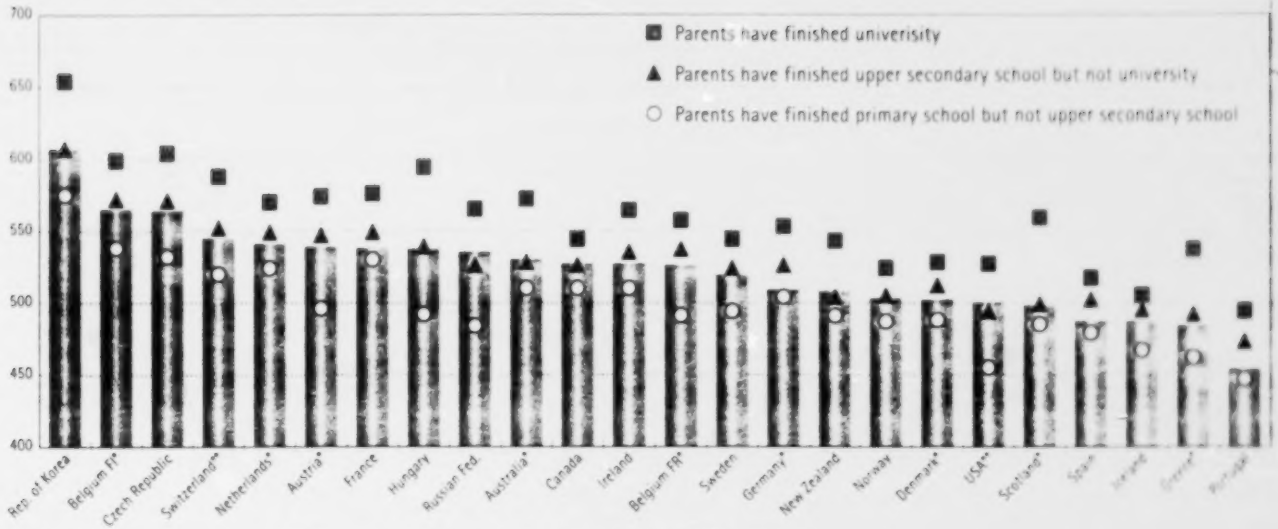
Data show that, in addition to wide overall fluctuations in achievement levels between various regions of a country, there are usually important disparities among different groups of pupils and between various types of schools. The most frequently cited disparities involve the following:

Socio-economic status – Pupils who come from homes with high socio-economic status (SES) as measured by factors such as family income, parental education and books in the home consistently score better on measures of achievement than pupils from low-SES families. This is a pattern that applies to countries of all kinds, including developed nations that have taken steps to guarantee equal educational opportunities for all.

Rural school, France © D'Si AMF's Agence VII

Figure 16

Mean mathematics achievement for 8th-grade students by parents' highest level of educational attainment (TIMSS)



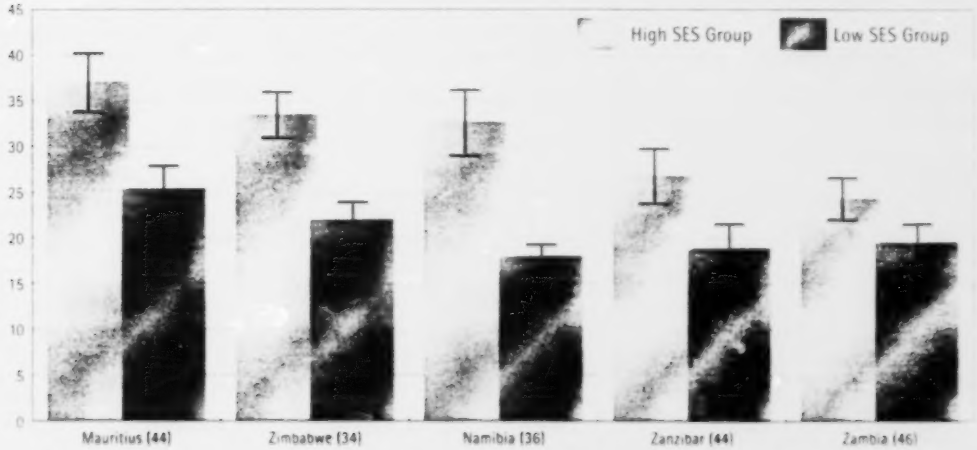
* Countries that do not follow TIMSS guidelines. ** Countries that follow partially TIMSS guidelines.
Source: International Association for the Evaluation of Educational Achievement (IEA)/ TIMSS, 1994-95.

In Figure 16 the bar indicates the mean achievement of Grade 8 pupils in twenty-four developed countries on the mathematical section of TIMSS, while the three symbols show the mean achievement levels of students with different levels of parental education. From this figure it can be seen that in all countries parental education continues to be an important source of disparities in pupil achievement.

Figure 17 shows differences in mean reading scores for the studies of Grade 6 pupils carried out by SACMEQ in southern Africa. The data show that pupils in the high SES group consistently outperformed those in the low SES group.

Figure 17

Reading mean scores for Grade 6 pupils from high and low socio-economic groups in five SACMEQ countries



Note: The figures in parenthesis indicate the number of "essential items" used in the tests.
Source: SACMEQ Policy Research: Reports No. 1-5. IIEP. UNESCO. 1998.

SACMEQ researchers also classified pupils into six groups using possessions in the home as a proxy for SES level and then calculated the proportion of pupils in each group who reached the minimum and desirable levels of reading mastery. Figures 18 and 19 show that while performance levels varied considerably from country to country, there was a consistent pattern of performance declining as SES moved from high to low.

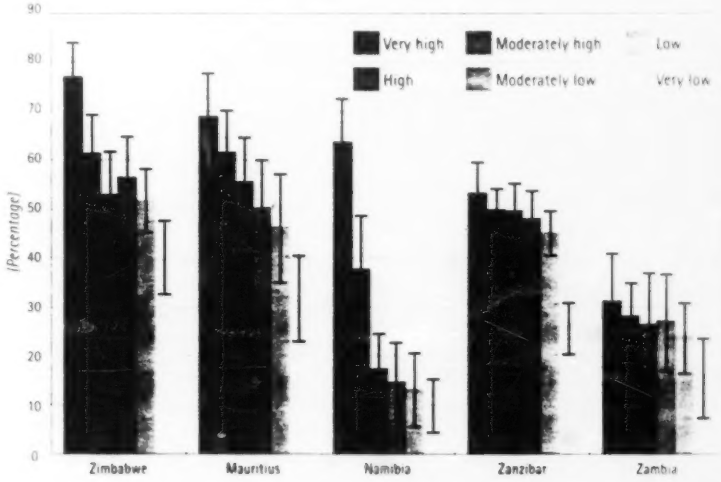
Gender - Data from developing countries frequently show that boys have an advantage over girls in areas relating to access to education, including retention and drop-out rates. The reasons for this advantage are complex and include the need for girls to work inside or outside the home, early marriage and pregnancy, less ambitious expectations by parents and a host of other cultural attitudes toward the education of girls and women.

However, patterns regarding gender differences in achievement vary widely not only from country to country but within countries depending on the subject matter, grade level and the types of pupils and schools being examined. In many, if not most, instances the variation in ability and achievement between the sexes is small when compared with the variation within groups of the same sex. This is confirmed by the most recent MLA studies in eleven African countries (Figure 20).

A PASEC study in the Côte d'Ivoire found that in Grade 2 boys performed at a higher rate in mathematics and girls in French, but both gaps narrowed as pupils approached the primary school exit point of Grade 5. The Zimbabwe study of reading literacy among Grade 6 pupils (Figure 6) found no important differences in the performances of boys and girls except for the rate of "illiterates", where boys had a statistically significant edge over girls. In Mozambique, MLA researchers found that boys did better than girls, especially in rural areas.[27] Data from the Laboratorio show that, with the exception of Paraguay, girls perform consistently better than boys in language. In Cuba, the Dominican Republic and Honduras, they also perform better than boys in mathematics.

Figure 18

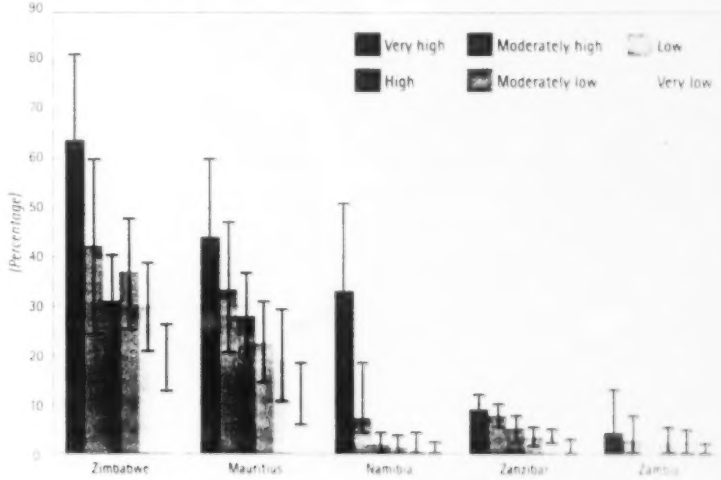
Percentage of Grade 6 pupils reaching minimum reading mastery level for different socio-economic level



Source: SACMEQ Policy Research: Reports No. 1-5. IIEP. UNESCO. 1998

Figure 19

Percentage of Grade 6 pupils reaching desirable reading mastery level for different socio-economic level



Source: SACMEQ Policy Research: Reports No. 1-5. IIEP. UNESCO. 1998

BEST COPY AVAILABLE

27. Education Update, p. 6.

GENDER

The complexity of the issue of gender differences in pupil performance is illustrated in a study carried out by Mioko Saito of Grade 6 reading literacy in Mauritius, Namibia, Zambia, Zanzibar and Zimbabwe. The study, sponsored by the Southern Africa Consortium for Monitoring Educational Quality (SACMEQ), measured the narrative, expository and document domains.

Saito concluded that there were 'no differences between boys and girls in terms of mean reading scores in the three domains of literacy and on total reading literacy score'. The author added, however, that 'the overall level of reading performance was generally very poor in all five countries.'

The central finding of the SACMEQ study is interesting because it contradicts evidence from studies in Mozambique and elsewhere showing that boys are doing much better than girls in examinations at the lower secondary level, and it conflicts with widespread evidence that illiteracy rates are higher among female adults.

On the other hand, Saito's conclusion is consistent with a 1991 national study of Grade 6 pupils in Zimbabwe. Such inconsistency in the findings of various studies, the author said, 'deserves further investigation'. One possible answer may of course be that fewer girls than boys enter and stay in school, thus the lower literacy rates among women.

Perhaps the most significant finding of the SACMEQ study relates to the large performance differentials that were documented in all five countries between the highest and lowest socio-economic groups, and between pupils in schools in cities and those in schools in rural areas. 'It appeared,' Saito concluded, 'that at the primary level, the differences between regions, socio-economic groups and school locations were much more important than gender differences.' This suggests the need not only for 'better allocation of teachers and resources' but 'a massive reform of the system'.

System reform may be needed also in countries where virtually all children attend school. According to a recent report from France* where most children are enrolled in pre-primary education at the age of 3, girls are doing better than boys at all levels. This does not, however, translate in equality in the job market, where high-level jobs are rarely held by women. Women are mainly making a career in service industries, where 60 per cent work in six occupations that represent only 30 per cent of the total job market. What distinguishes girls and boys at school is the choice of courses. Science courses are dominated by boys, social, economic and literary courses by girls. Since high-level technical and scientific skills are often required for high-level jobs, gender inequality in the work place can be partly explained by previous academic choices.

But why do girls and boys still make such different choices at school? Researchers have shown that children adapt to gender roles according to expectations from society at large. Gender roles are learned by observation and imitation in the family, in the media and at school. School transmits knowledge, but it also transmits norms, values and social models. There, pupils learn to take an interest in subjects as a function of gender-specific expectations they have acquired.

Thus, through very subtle processes, mostly unconscious both to pupils and teachers, schools treat girls and boys differently. To fight this sort of inequality, assessments of both formal and socio-psychological learning have to be conducted. Schools, instead of contributing to the reproduction of gender inequalities, could become a powerful tool for equality and emancipation. This is not a problem of girls but a political and social question that concern girls and boys, women and men.

* F. Vouillot (ed.), *Filles et Garçons: une egalité a construire*, Paris, CNDP, 1999.

In an analysis of results from the SACMEQ studies of Grade 6 reading literacy in southern Africa, Mioko Saito found no significant differences between boys and girls in reading. She noted the sometimes contradictory findings of various studies and suggested that the issue of gender differentials warrants further study. Saito added, however, that whatever these differentials may be, their significance pales in comparison to differentials of regions, school location and socio-economic status.[28]

A number of studies over the years have found that female primary and lower secondary school pupils do better than males in language-related subjects, while the reverse is true for mathematics and science. Findings from more recent studies, however, show that these gender differences that are clear in the first years of schooling have a tendency to become reduced as pupils move to higher grades.

28. Mioko, Saito, *Gender vs. Socio-Economic Status and School Location Differences in Grade 6 Reading Literacy in Five African Countries*, SACMEQ Report 98.001, p. 259, UNESCO, Harare, Zimbabwe.

BEST COPY AVAILABLE

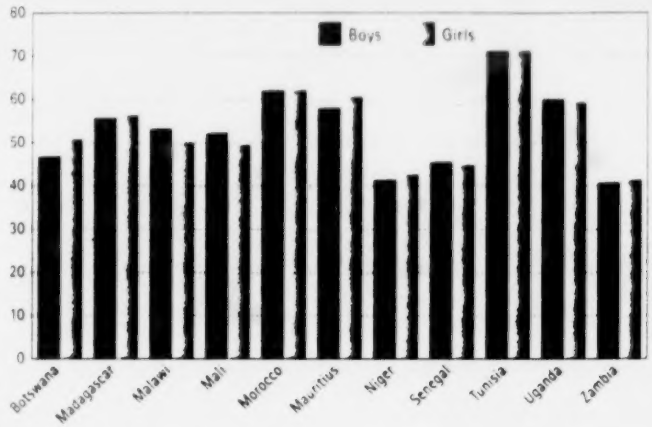
The IEA Reading Literacy Study, however, undertaken in 1990 and 1991 reported that girls outperformed boys in all participating countries. In the same vein, the IEA's first international science study found not only that boys outperformed girls in the sciences but that the gender gap widened as pupils moved to higher grades. Sex was found to be a weak predictor of science achievement in seventeen countries, three of them developing nations[29]. Gender contributed only two per cent of the total explained variance in science achievement, though the impact was much more significant in the developed countries. The Second International Science Survey found that boys outscored girls at all levels.

No common pattern emerged in gender differences from TIMSS. Nevertheless, while gender differences among students in Grades 7 and 8 were relatively low in mathematics, those in science were more pronounced. Boys performed better in half of the participating countries, especially in Grade 8.

School Location – Pupils in urban areas perform, on average, better than their counterparts in rural schools. The reasons generally given include the fact that big cities and, to a lesser extent, mid-sized urban areas have relatively large proportions of high-SES families. Schools in such areas often have better facilities and are in a favourable position to attract good teachers.

Figure 20

Global mean score by gender in eleven African MLA countries



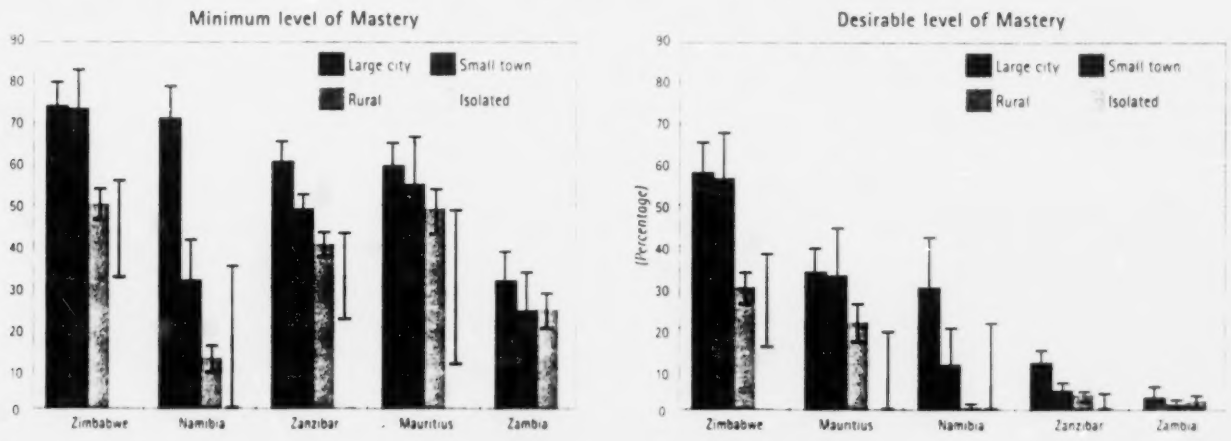
Source: V. Chinapan, et al. *With Africa for Africa: Towards Quality Education for All. Draft Regional Report.* EFA 2000 Assessment. MLA Project, 1999.

The relation between location and achievement can be seen vividly in Figure 21, which shows the relative achievement of pupils in large cities, small towns, rural areas and isolated venues. With some exceptions – pupils in isolated areas of Namibia did better than their rural counterparts – the findings tend to show a clear pattern of declining achievement as the setting moves from high to low population density.

29. Comber, L.C. and Keeves, J.P., *Science Education in Nineteen countries. An Empirical study.* Stockholm: Almqvist & Wiksell, 1973.

Figure 21

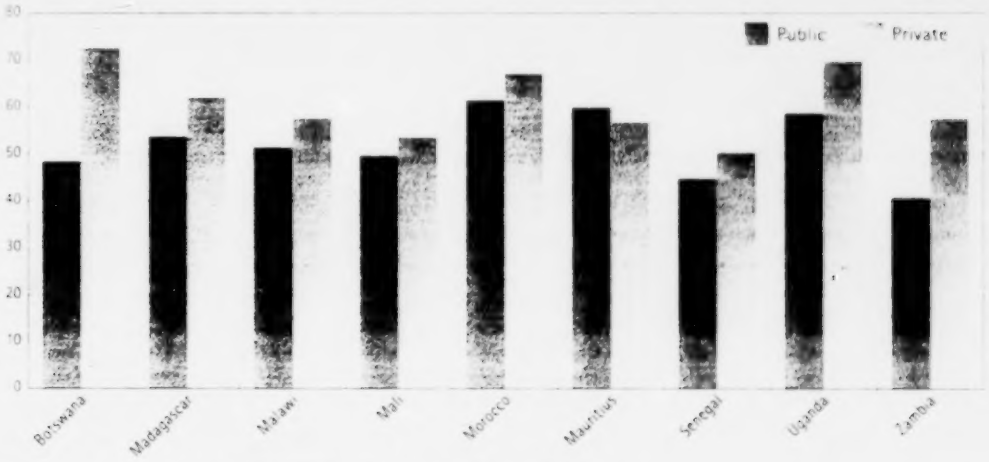
Percentage of pupils reaching minimum and desirable mastery levels for different school locations



Note: For Zambia, Isolated area is included with Rural.
Source: SACMEQ Policy Research: Reports No. 1-5. IIEP. UNESCO: 1998.

Figure 22

Global mean score by type of school in eleven African MLA countries



Source: V. Chinapah, et al. *With Africa for Africa: Towards Quality Education for All. Draft Regional Report. EFA 2000 Assessment*. MLA Project, 1999.

There are, however, many exceptions to generalizations about school location. The Laboratorio study of thirteen Latin American school systems found that some rural schools in Colombia outperformed schools in both large and mid-sized urban areas.[30] Significantly, Colombia has mounted initiatives aimed at improving rural schools.

30. *Latin American Laboratory for the Assessment of Quality in Education. First International Comparative Study*. UNESCO Santiago, Chile, 1998.

Types of schools - Most national assessment data show that, on average, pupils in private schools tend to perform at a higher level than do those in public schools. Explanations usually include the fact that parents who send their children to private schools tend to have higher incomes and educational backgrounds, and that private schools often have better

Table 2

"Within-country differences" in Basic Learning Competencies (BLCs) mean scores by region (urban/rural), gender and school-type (public/private) in ten African MLA countries

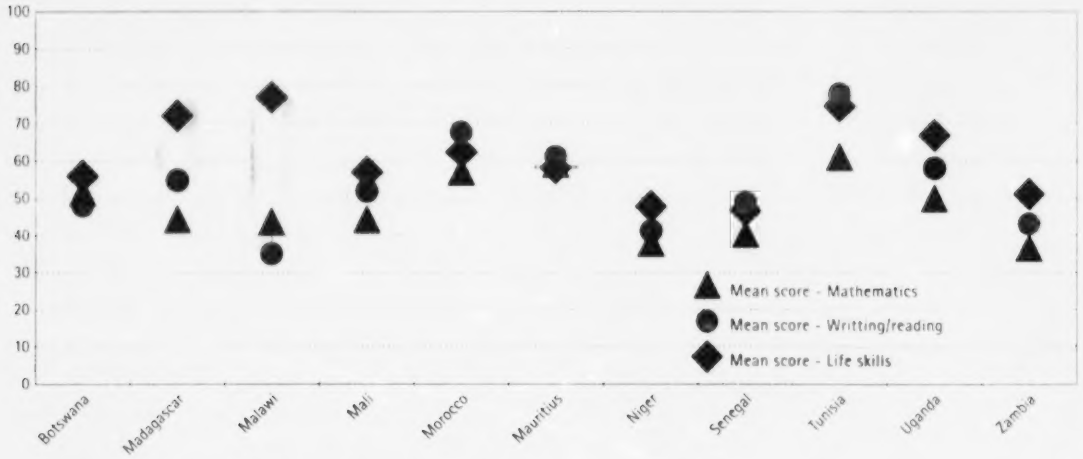
	Botswana	Madagascar	Malawi	Mali	Morocco	Mauritius	Niger	Senegal	Tunisia	Uganda
Life Skills										
Grade IV										
Mean Point-Score Differences										
Urban vs. Rural	0	4	2	11	5	4	0	0	7	1
Girls vs. Boys	**3**	**1**	**-4**	**-2**	**0**	**0**	**0**	**0**	**0**	**-3**
Private vs. Public	19	4	5	8	5	-1		-1		10
Literacy										
Grade IV										
Mean Point-Score Differences										
Urban vs. Rural	4	13	2	7	3	6	5	2	6	7
Girls vs. Boys	**5**	**2**	**-1**	**-5**	**-1**	**5**	**1**	**-1**	**0**	**1**
Private vs. Public	-35	15	9	3	4	-4		1		16
Numeracy										
Grade IV										
Mean Point-Score Differences										
Urban vs. Rural	2	11	2	2	5	7	2	5	9	4
Girls vs. Boys	**3**	**-1**	**-5**	**-2**	**1**	**2**	**1**	**-1**	**-2**	**-3**
Private vs. Public	22	8	6	1	7	-4		15		8

Source: Chinapah, V. et al. With Africa for Africa: Towards Quality Education for All. Draft Regional Report. EFA 2000 Assessment. MLA Project, 1999.

BEST COPY AVAILABLE

Figure 23

Mean scores by subject in eleven MLA African countries



Source: V. Chinapah, et al. *With Africa for Africa: Towards Quality Education for All. Draft Regional Report. EFA 2000 Assessment.* MLA Project, 1999.

physical and human resources. Figure 22 shows that pupils in private schools outperformed those in public schools in ten out of eleven MLA African countries, Mauritius being the exception.

The MLA researchers in Mali looked not only at public and private schools but at community-based fee-paying schools and at *medersas*, (privately-supported schools that use Arabic as the medium of instruction). They found that the public schools and medersas were the least efficient. Similarly, the MLA study of Jordan found that private school pupils outperformed those in schools run by the Ministry of Education in all four subjects measured.

Once again, however, such performance patterns are by no means universal. The Laboratorio research found exceptions in Bolivia, Venezuela and Honduras.[31]

A good overview of the extent to which achievement typically varies among different groups of pupils and types of schools can be seen in Table 2 summarizing the results of the MLA studies of life skills, literacy and numeracy among Grade 4 pupils in 10 countries in Africa.[32]

These data confirm the general conclusion that urban pupils score higher than rural ones, although the margin of difference varies from modest in Bostvana and Malawi to quite dramatic in Madagascar. The data also show that in almost half of the cases, the mean scores for girls were higher than those of boys. In Mali and Malawi, boys did slightly better than girls in all three subjects. Comparisons of pupils in private and public schools show that in almost all cases the privately-schooled children did better. A conspicuous exception was Mauritius, where public-school students had higher scores in all three subject domains.

The importance of disaggregating data

Disparities in achievement among various groups of pupils raise important policy issues for national education systems, and analysis of the nature of these disparities can be an effective tool for addressing these issues. Such data can, among other things, identify students and schools with the greatest educational needs and guide decisions about how to allocate educational resources in the most effective way. They also highlight the importance of tailoring curriculum and teaching strategies to the differing needs of pupils. In order to make maximum use of results from national assessments, it is important to look beyond aggregated data.

31. *Latin American Laboratory for the Assessment of Quality in Education. First International Comparative Study.* UNESCO Santiago, Chile, 1998

32. Vinayagum Chinapah et al. *With Africa for Africa: Towards Quality Education for All. Draft Regional Report. EFA 2000 Assessment.* MLA Project, 1999.

BEST COPY AVAILABLE

Table 3

Mean scores by sub-domain in eleven MLA African countries

	Reading / Writting				Mathematics			Life skills		
	Vocabulary	Under-standing	Grammar	Writting	Arithmetics	Measuring	Geometry	Health	Civic life / Environment	Practical skills
Botswana	70	53	49	27	53	39	54	54	69	49
Madagascar	53	72	48	56	49	32	43	80	72	76
Malawi	54	37	35	23	42	43	47	78	80	70
Mali	85	57	44	45	34	50	48	56	58	56
Morocco	77	99	66	58	47	56	71	62	65	57
Mauritius	86	68	57	49	61	52	61	56	66	55
Niger	81	51	43	33	36	43	46	46	49	49
Senegal	75	48	49	36	29	39	38	48	48	41
Tunisia	75	81	76	78	63	55	70	70	80	72
Uganda	80	60	59	45	49	48	53	68	66	64
Zambia	72	45	49	28	36	35	37	52	52	48

Source: V. Chinapah, et al. *With Africa for Africa: Towards Quality Education for All. Draft Regional Report. EFA 2000 Assessment.* MLA Project, 1999

The five initial MLA studies offer a good example of how in-depth analysis can yield important insights into learning patterns in particular countries.

Analysis of the performance data from China found that in the life-skills area pupils performed better in health, nutrition, safety and everyday life than they do in the labour sub-domain. In the area of literacy most pupils were found to have acquired solid basic knowledge and proficiency, but they tended to be weak in the areas of reading and writing. Similarly, in numeracy the majority of pupils were shown to have mastered basic calculating skills, but they were weak in statistics and in solving practical problems.

As shown in Figure 23, recent MLA studies found that in most countries, the highest scores were obtained in life skills and the lowest in mathematics. Table 3 below shows the mean scores in different sub-domains within the three main subjects (life skills, writing/reading and mathematics)

Such insights suggested a number of new policy directions in China, including more training in life skills at the early childhood stage and the tuning of the mathematics syllabus to take greater account of specific sub-domains. The analysis points to shortcomings of the conventional educational practice that 'puts emphasis on knowledge but neglects practical skills.'[33]

As already suggested, data from the Mali MLA study showed that pupils in private schools outscored those from public schools in numeracy. When results are broken down into various sub-domains, however, different tendencies are observed. Pupils in *medersas* outscored all other pupils in the basic operations of subtraction, multiplication and division, while the opposite occurred in the more sophisticated skills of geometry and arithmetic. Such insights offer potentially valuable guidance for educators seeking to implement child-centered teaching strategies. "However, as Vinayagum Chinapah observed, such strategies 'will remain perpetual dreams' as long as the information base for monitoring is crude and incomplete."[34] ■

BEST COPY AVAILABLE

33. V Chinapah, *Handbook on Monitoring Learning Achiement: Towards Capacity Building*, UNESCIO, 1997.

34. Ibid, pp. 84-5.

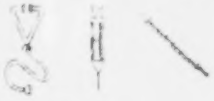
ASSESSING LIFE SKILLS

MLA tests

In addition to literacy and numeracy tests, the Monitoring Learning Achievement Project assesses the mastery of life skills. Prototype questions developed by UNESCO experts serve as a basis for countries to develop their own tests adapted to their particular socio-cultural contexts. The result is both a common core of questions that all countries can use as well as a set of specific questions based on country-specific criteria. Following are some examples:

Health/hygiene/nutrition

Your little sister is ill, you touch her forehead and it is very hot. Which of the following instruments would you use to see if she has fever?

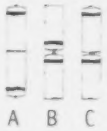


Before eating, you must always:

❑ a. *wash your face*
❑ b. *wash your hands*
❑ c. *comb your hair*
❑ d. *I don't know*

Everyday life

Which is the right way to join two batteries when you change batteries for flashlights?



❑ A. ❑ B. ❑ C. ❑ D. *no idea*

You are walking with a friend.
Suddenly a storm breaks out. What do you do?

❑ a. *you wait under a three*
❑ b. *you go home*
❑ c. *you continue walking*
❑ d. *you avoid walking under electrical lines*

Social/natural environment

Is it right or not to put urine dirty water into the river nearby?

❑ a. *right. It is very convenient*
❑ b. *wrong. It will pollute the river*
❑ c. *right or wrong, neighbours do so, too*
❑ d. *no idea*

You find a watch on the ground near school. What do you do?

❑ a. *I give to the teacher*
❑ b. *I keep it for myself*
❑ c. *I keep it to sell*
❑ d. *I don't know.*

Source: Handbook on Monitoring Learning Achievement Towards Capacity Building, by Vinayagum Chinapah, UNESCO 1997.

The International Life Skills Survey

The International Adult Literacy Survey has shown that direct measures of skills are better predictors of successful participation in society than previous indicators such as education credentials. The International Life Skills Survey (IALS), now under development by OECD, Statistics Canada and the National Center for Education Statistics in the United States, measures a wide array of basic life skills directly linked to social and economic success of the adult population. In addition to the traditional areas of Prose and Document Literacy and Numeracy, IALS measures:

Problem Solving

The ability to successfully find solutions for problem situations with no previous knowledge of the solution path is increasingly important is today's complex world. The study will provide policy makers, employers and educational planners with data about levels of problem-solving ability.

Practical Cognition

Sometimes referred to as Practical Intelligence, Tacit Knowledge or Common Sense, this area refers to knowledge relevant to problems that are not clearly defined, personal in nature, relate to everyday experience, have multiple "correct solutions" and have multiple methods for deriving them. Measurement relies on items that elicit responses to a real life situation.

Teamwork

Governments, businesses and community groups rely increasingly on work teams to streamline processes, enhance participation and improve performance. Information from assessment of the interpersonal and other skills required by teamwork can be used by employers and educators to assess and improve teamwork in a nation's workforce and general population.

Information and Communication Technology Literacy

This area describes the ability of people to make full use of existing, new and emerging technology in order to be successful both in professional and private life. Measurement involves extended scenarios in which common activities are described and subjects are asked how to change the scenario by making "good use" of information and communication technology.

BEST COPY AVAILABLE

PART V

Factors that contribute to achievement

BEST COPY AVAILABLE

Most of the findings presented in the previous sections attempted to answer questions related to pupil achievement levels and to disparities in different countries. Discussions of disparity issues leads directly to policy questions such as: What is the relative impact on learning of educational policies vis-à-vis environmental factors? What educational inputs have the most impact on pupil achievement? What groups of pupils are most likely to fail and thus might benefit from targeted interventions?

The analyses described thus far do not allow us to answer such questions properly because they are simply based on comparisons of means. For instance, data may show that pupils in urban areas outperformed those in rural areas, but they will not tell us whether this disparity is due to the geographical location or whether it can be explained by related conditions, such as the fact that many pupils in rural areas come from poorer families with less educated parents or that their schools have limited educational resources. These categories overlap one another, of course, and they do not imply any causal effect. However, they can be useful in identifying some populations to target in the interest of educational equity. In order to assess the causal effect of single factors on learning achievement, one needs to control for the simultaneous effects of other factors, by applying multivariate statistical techniques.

An enormous literature exists on the question of which inputs and combinations of factors will produce the greatest educational outputs as measured by student achievement. This report is not the place to review this literature, especially as it relates to generalizations that apply across national borders. Nevertheless, many of the national assessments under review have identified correlations between various inputs and improved student learning in their particular conditions, and analysts have used these correlations to suggest policy changes aimed at enhancing student achievement.

Almost all research findings have shown that we cannot identify a unique determinant of pupils achievement, be it teacher qualifications, the availability of learning materials or the socio-economic status of the pupils' families. In thinking about such matters, researchers have found it helpful to divide factors into two broad categories: contextual and school-related.

Contextual factors that impact on achievement are related to the context in which schools and individual students function. These include elements such as whether the school is in an urban or a rural area, the socio-economic level of the community and the educational attainment of parents. State education systems face the challenge of providing equal learning opportunities to pupils coming from a diversity of socio-economic, cultural, attitudinal and other contexts. By and large, school officials have little control over contextual factors.

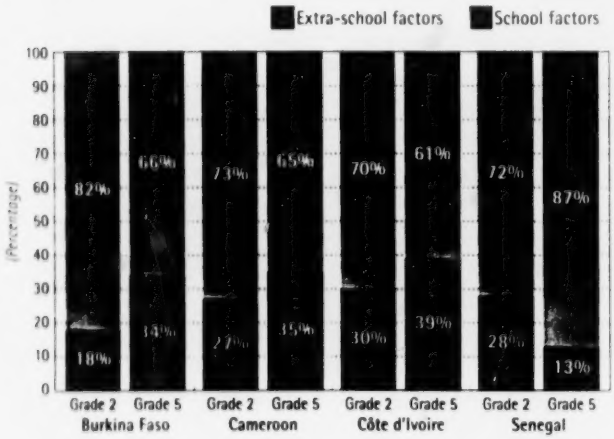
School-related factors are a function of school policies defined by political and educational leaders at the national, district or local level. These include elements such as retention policies, the qualifications of teachers, the length of the school year and day, homework policies, the availability of textbooks and other educational materials, and how convenient schools are to where pupils live. By definition, educational administrators and policy-makers have considerable influence over school-related factors.

BEST COPY AVAILABLE

Medical exam, Senegal. © *Thien LAM DUC/Agence VU*

Figure 24

Distribution of the explained learning gain between school factors and extra-school factors



Source: Programme d'Analyse des Systèmes Educatifs (PASEC). Les facteurs de l'efficacité de l'enseignement primaire: données et résultats sur cinq pays d'Afrique et de l'Océan Indien. CD-ROM (Version de mars 1999).

The *in-school factors* known to influence achievement can be grouped into teacher characteristics and school resources. Teacher characteristics are of paramount interest because the qualification, experience and competence of the teachers play a critical role in the shaping the process of teaching and learning and because the interactions between pupil and teacher are the primary way of transmitting knowledge and skills. School resources consist of factors such as facilities, the availability of textbooks, class organisation and instructional procedures.

Relative impact of contextual and school-related factors

While both contextual and school-related factors are present and significant in all countries, their relative influence varies from country to country. In general, contextual factors are particularly useful in explaining achievement differentials among pupils in developed countries, whereas school-related factors tend to be relatively important determinants of such differences in developing nations

Studies in industrial nations of Europe and North America have consistently found a negative correlation between family size and educational achievement, while similar studies in Kenya and the United Republic of Tanzania have found that the correlation is positive. By contrast, the availability of teaching materials is strongly correlated with student performance in developing countries but not in developed countries.[35] Many other studies seem to show that the influence of the home environment on scholastic

35. Greaney and Kellaghan, 1996, op. cit., p. 36.

performance is weaker in developing countries than in developed countries.

The PASEC study analysed the variance of the learning gains between pupils in four countries and found that the school-related factors explain 61 to 87 per cent of the explained score differences, while pupils' families and their socio-economic background generate only 13 to 39 per cent of these differences (Figure 24). The PASEC multivariate regressions were based on eight contextual and sixteen in-school variables.

Findings from several earlier researches contrast with those presented by PASEC. A 1980 review of studies dealing with determinants of pupil achievement in developing countries concluded that many inputs, including class-size, school facilities, per pupil expenditure, instructional time and innovation in instruction methods, did not contribute substantially to gains in scholastic achievement and that pupil socio-economic status is consistently the major determinant of academic achievement.[36] According to Husén, this does not mean that "schooling does not make any difference".[37] Rather, it simply suggests that above a certain threshold of resources, where most schools in so-called developed countries are located, home background contributes more to differential learning outcomes than schooling.

Effects of contextual factors on achievement

The personal characteristics and family environments of pupils are go far to explain variations in learning achievement. Findings from international surveys have pointed out the impact of factors ranging from parental education and expectations to family size and the distance to school.

As seen in Figure 25, PASEC researchers looked at five such factors and found robust impacts on learning mathematics and French. They found that the opportunity to speak French at home is a real advantage, particularly in Grade 2. The data showed that rural areas are unfavourable to high achievement levels – an effect that is probably not due to the standards of living or the cultural level of the parents (both of which have been controlled for) but rather to geographical and more general socio-cultural reasons. Children who were over-aged at their arrival at school were handicapped for the rest of their school life. This suggests that authorities in rural areas, where compulsory attendance laws may not be observed in a timely fashion, should inform parents about the importance of sending their children to school on time.

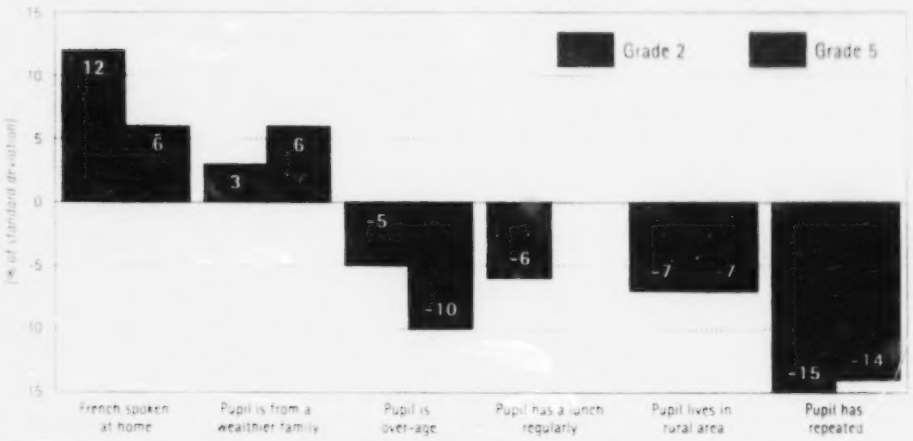
Primary school, France. © Gilles LARVOR/Agence VU

36. J. Simmons, and L. Alexander, Factors which Promote School Achievement in Developing Countries: A Review of the Research. In J. Simmons (ed.) *The Education Dilemma: Policy Issues for Developing Countries in the 1980s*. Oxford: Pergamon Press, 1980.

37. T. Husen, *Education and the Global Concern*. Oxford: Pergamon Press, 1990.

Figure 25

Causal effects of the extra-school factors on mathematics and French achievement in five PASEC countries.



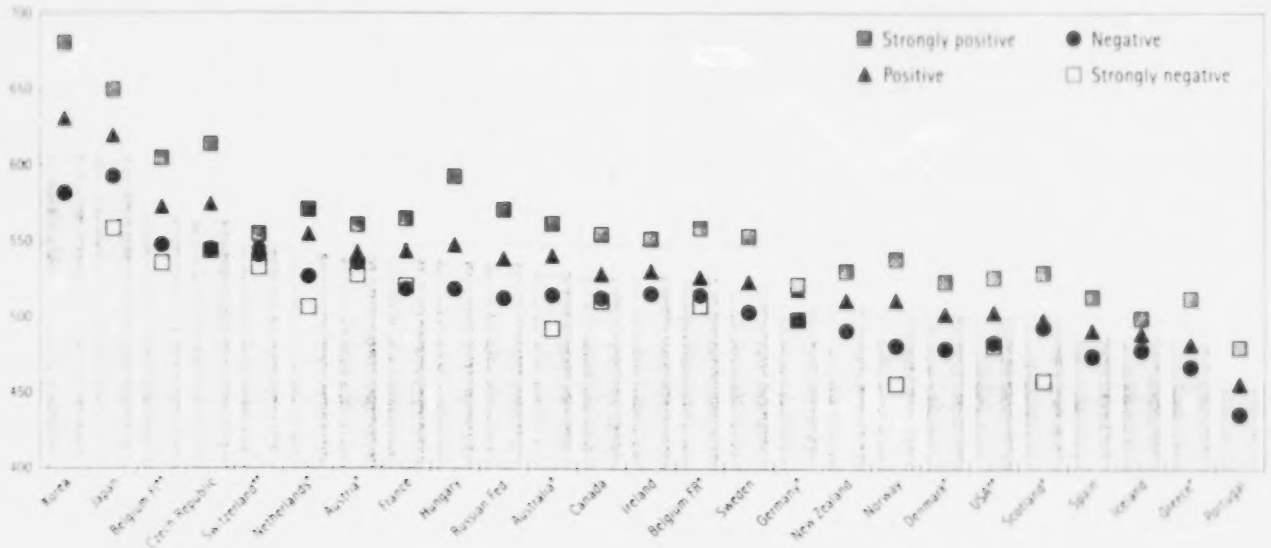
Source: Programme d'Analyse des Systemes Educatifs (PASEC). *Les facteurs de l'efficacite de l'enseignement primaire: donnees et resultats sur cinq pays d'Afrique et de l'Ocean Indien.* CD-ROM (Mars 1999).

Pupil personal factors such as attitude and motivation have been found to influence learning in a wide variety of subjects under various conditions. In fact, pupils' perceptions of the value of learning particular subjects may be considered as both inputs and outcomes of educational process, since their attitudes can be related to educational achievement in ways that reinforce higher or lower performance.

Figure 26 shows how the attitudes of Grade 8 pupils toward mathematics correlated with their performance on the TIMSS test. The bar shows the mean achievement among all pupils and the three symbols indicate the mean mathematics achievement of pupils with different attitudes towards mathematics. The data demonstrate a consistent pattern of pupils who do well in mathematics having a generally more positive attitude towards this subject than do other test takers.

Figure 26

Mean mathematics achievement scores of Grade 8 pupils and their attitudes towards mathematics (TIMSS)



* Countries that do not follow TIMSS guidelines. ** Countries that follow partially TIMSS guidelines.
Source: International Association for the Evaluation of Educational Achievement (IEA)/TIMSS, 1994-95.

MLA researchers in Jordan analysed questionnaires administered to pupils, parents and educators, and found, as might be expected, a strong correlation between parental education and the success of children in Arabic, mathematics and science. They also found that pupils who help out in their homes for one hour or less achieved at a higher rate than those who work for two or more hours. While such data might at first glance appear to make a strong case for limiting the amount of work that children do – a significant issue in countries that lack child labour laws – researchers in Jordan concluded that such an interpretation was not appropriate to their country. The correlation between modest amounts of chores and high achievement, they said, was due to the fact that children from wealthier homes were not required to do as much work around the house as their less advantaged counterparts.[38]

MLA researchers in Morocco identified both contextual and school-related factors that impacted on pupil performance. The primary contextual influence was the amount of support pupils received in their homes. The more support and help that the child received at home, they found, the better his or her chances of scholastic success. In-depth analysis of the way various factors impacted on each other showed that parental support was important in minimizing absenteeism, which has a strong negative effect on achievement. One policy implication was that convincing parents to make sure that their children completed homework assignments was essential to educational success.[39]

Effects of school related-factors on achievement

Moroccan MLA researchers also identified a number of school-related factors that impacted on achievement. They found, for example, that academic performance was inversely correlated with the amount of time that pupils spent getting to and from school. Improving the organization of school transport thus emerged as a reasonable priority for policy-makers. They also developed data showing that an increase in the number of classrooms correlated with higher achievement and that pupils who are in the appropriate grade for their age tend to enjoy more academic success.

The policy implication of this finding was that forcing substantial numbers of pupils to repeat grades was neither academically sound nor cost-effective. Other school-related factors that were identified showed that the physical condition of schools and the background of teachers had an impact on pupil performance in numeracy and in Arabic but not in life skills.[40]

The District Primary Education Programme (DPEP) assessment in India found that 'districts having [the] maximum number of schools holding the complete range of competency based materials' registered higher pupil achievement than did other districts.[41]

In 1994, five years after it began assessing the quality of its education system, Mexico released a

Walking to a rural school, France. © B. BESCAMPS/Agence VU

38. Chinapah, op. cit. pp. 86-9.

39. Ibid., pp. 94-101.

40. Ibid., pp. 94-101.

41. DPEP Core Resource Group, National Council of Educational Research and Training, *Mid-Term Assessment Survey: An Appraisal of Students' Achievement*, New Delhi, October 1998.

report on the knowledge and skills of 480,000 teachers and the achievement of 2.8 million primary and secondary school children. Among other things, the research concluded that children who attended pre-school performed better than those who did not.[42]

The SACMEQ study in Mauritius found a high correlation between reading proficiency and going to extra tuition after school.[43] In a similar vein, the MLA study in Nigeria concluded that the low overall level of achievement reported was 'largely due to very low level of literacy competence on the part of the pupils (possibly and their teachers too)'. The report concluded, 'There is therefore an urgent need to improve the teaching of English Language.'[44] Likewise, the MLA survey in Mauritius found that pupils on the island of Rodrigues were having particular difficulty with English and that this lack of proficiency was affecting their performance in numeracy and life skills.[45]

A study by the Center for Educational Research and Development in Lebanon concluded that Grade 4 pupils were performing at unacceptably low levels and that the lowest achievement levels were found among pupils who, among other things, did not have pre-school experience, who had to travel more than an hour to get to school, who had paid jobs and whose teachers and principals lacked high levels of professional training.[46]

Figure 27 shows the impact of some school factors on the scores in mathematics and French for the Grade 2 and Grade 5 pupils in five PASEC countries. Researchers found that out of sixteen school factors used in the regression model, thirteen have a significant impact on learning achievement. The impacts, however, were by no means consistent. Having a woman as teacher and pre-service training had a positive impact on learning in Grade 2 but a negative impact at Grade 5. Likewise, repetition seems to have a negative effect in learning achievement, as did teacher on-job training.[47]

PASEC researchers looked in detail at three in-school variables:[48]

Textbooks. Researchers found positive impacts of textbooks in the five countries studied.

In general, French and mathematics textbooks have a positive impact on learning for children in Grade 2. The French textbook appears to be the most important, with the maths textbook bringing only a small and statistically insignificant benefit. French textbooks have a positive impact equivalent to the gain of two ranks in a class of 100 pupils (or 7 points of standard error on scores).

Results are less clear in Grade 5. Statistically positive effects are reported only for access to both French and mathematics textbooks in Senegal and Burkina Faso. In other countries, and for the French textbook alone, the estimated effects are too weak to be statistically significant. Thus it seems that textbooks are more efficient at the beginning of primary education. One puzzling finding is the absence of any visible impact of textbooks in Côte d'Ivoire.

Teacher recruitment and training. This issue can be divided into three variables: the academic level of recruitment, pre-service training and on-the-job training. The data are decisive but difficult to evaluate, in part because the sample included only 120 teachers in each grade in each country. Nevertheless, analysis leads to some results that warrant further research.

Recruitment of teachers was studied at three academic levels: tenth grade or less, more than tenth grade but without graduation from secondary school, and graduation from secondary school or more. In general, there is a small gain in going from the first to the second category, but going to the third category produced no gain or even to a loss. Thus there seems to be a weak, and statistically non-significant, benefit from additional years of academic studies by teachers, but this gain is quite soon saturated.

Analysis of the impact of pre-service training is complicated by the fact that it was not possible to control for other dimensions of such training, such as its duration or content. Moreover, results lack consistency from one country to another. Nevertheless, it seems that pre-service training has a positive impact of the order of 4 per cent of standard deviation a year. At the same time, the data suggest some problems. In Grade 5 teachers who benefited from one year of training in Burkina Faso or Cameroon have lower results than those who received no training. The same pattern holds

42. Laurence Wolff, 'Educational Assessments in Latin America: Current Progress and Future Challenges,' Occasional Paper Series No. 11, Partnership for Educational Revitalization in the Americas, Washington DC, p. 10.

43. Kulpool, op. cit., p. 77.

44. Wole Falayajo, Gladys A. E. Makoju, Peter Okebukola, David C. Onuqha and Joseph O. Olubodun, *Assessment of Learning Achievement of Primary Four Pupils in Nigeria*, MLA Project, September 1997.

45. *A Survey of 9 Year Old Children in Rodriguan Schools in Literacy, Numeracy and Life Skills*, Mauritius Examinations Syndicate, April, 1996.

46. Center for Educational Research and Development, Lebanon, 1995.

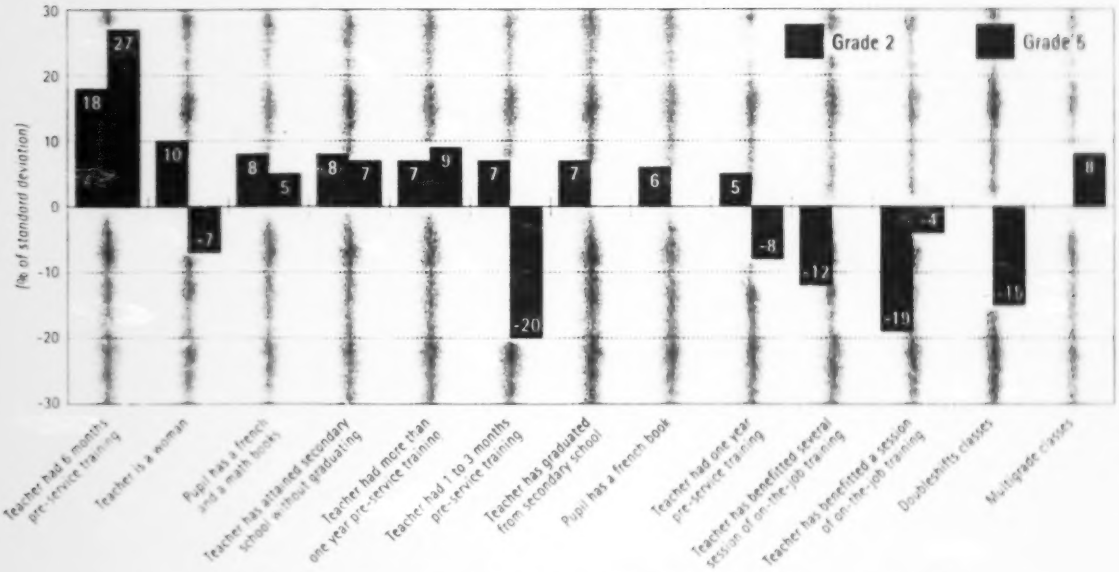
47. The PASEC used standardized scores (or Z-score) to facilitate comparisons in achievement between countries and between other sub-groups of pupils. The Z-score is a mathematical transformation of raw scores taking into account the mean and the standard deviation of the total distribution for the purpose of referencing an individual's score to the total distribution of scores. The distribution of standard scores has a mean of **0** and a standard deviation of **1**. Hence the causal effects (coefficient of the regression) presented in this graph can be interpreted in terms of units of standard deviation.

48. Programme d'Analyse des Systèmes Educatifs (PASEC), Les facteurs de l'efficacite de l'enseignement primaire: données et résultats sur cinq pays d'Afrique et de l'Ocean Indien, CD-ROM (Version de mars 1999).

BEST COPY AVAILABLE

Figure 27

Causal effects of the school factors on mathematics and French achievement in five PASEC countries



Source: Programme d'Analyse des Systèmes Educatifs (PASEC). Les facteurs de l'efficacité de l'enseignement primaire: données et résultats sur cinq pays d'Afrique et de l'Océan Indien. CD-ROM (Mars 1999).

in Senegal for teachers who received two years of training rather than one. Thus, it seems that some cases of training are truly inefficient. Further research would be particularly welcome on this point.

Data on the impact of on-the-job training in the four countries show that, with the exception of Grade 2 in Senegal, one period of training has no positive – and even a negative – effect, but several periods have a strong positive impact. This finding suggests that training pays if it is repeated.

Class organization. Data on this topic cover three variables: multi-grade classes, double shifts and class size. Fortunately, results in these areas are consistent across grades and across countries. They show that multi-grade and double shift classes impede pupil learning, as do large classes. The question is: at which scale?

Multigrade classes gather pupils from different grades around one teacher. Except in Senegal, the results show a loss of around 15 per cent of standard error in Grade 2 (the loss is not significant in Grade 5, where the point estimate is -5 per cent). However, the results in Senegal reported a positive impact, especially in Grade 5. In a comparable study for Togo, J. P. Jarousse and A. Mingat also documented a strong positive effect of multi-grade classes and documented the positive impact of this pedagogical method as such.

Double-shift classes make it possible to have two student cohorts working with the same teacher in the same classroom at different times in the day, typically morning and afternoon. Data show that the general effect is to lower children's learning, although Cameroon in Grade 5 is a curious exception. Generally speaking, the loss amounts to 7 per cent of standard deviation in Grade 2 and 16 per cent of standard deviation in Grade 5.

Larger classes seem to have a very weak negative effect on children's learning: almost nothing in Grade 2 and -2 per cent of standard deviation in Grade 5 for ten additional children in the class. The striking fact, therefore, is the quasi-absence of effect of larger classes. This might be explained by the relatively authoritarian pedagogy in the countries studied, which does not lend itself to reaping the potential advantages of smaller classes through more personal teaching relationships. ■

BEST COPY AVAILABLE

PART VI

How countries use assessment to improve student learning

Primary school, Paris, France. Agnes BONNOT/Agence VU

BEST COPY AVAILABLE

Developing countries have used information gathered through national assessments to improve teaching and learning in a wide variety of ways. This section will discuss some of these approaches.

Targeting of scarce resources

A number of developing countries have made effective use of assessment data to identify significant learning problems and then to direct financial and other resources toward addressing them. Such targeting allows these countries to make more effective and efficient use of scarce resources.

Chile established a national assessment system in 1978, but in 1984 a new minister of education abolished the programme, ostensibly for reasons of cost. However, in 1988, following widespread decentralization of national institutions, the assessment programme was revived and reorganized as the Sistema de Medición de Calidad de la Educación (SIMCE). The programme tests all pupils in Grades 4 and 8 in Spanish and arithmetic, and a 10 per cent sample of pupils is tested in the natural sciences, history and geography. Information is also obtained on attitudes and backgrounds of teachers and parents, and on school efficiency; teachers received reports of how their classes did on each of the assessed objectives.[49]

Laurence Wolff of the Inter-American Development Bank described the SIMCE programme as 'the most comprehensive and best-managed assessment system in Latin America,' adding that it has "served as a strong tool for implementing a reform program fostering decentralization, accountability and increased learning."[50]

SIMCE scores were used to identify 900 – and subsequently 1,200 – of Chile's poorest-performing schools. These schools then became the focal point of a pedagogical programme that channeled additional books and other educational materials into these schools and provided them with in-service training of teachers and other forms of infrastructure support. Financial rewards are also offered to schools where test scores increase from yearly. By such a measure, the programme has been judged a success.

Shaping classroom teaching

Since 'teaching to the test' is a common phenomenon, at least with public examinations where success or failure has important consequences, measurement devices can be used to influence the topics that teachers emphasize in their teaching. Thailand organized a national assessment programme in 1981 at the same time that it introduced a new higher secondary school curriculum in 1981. The assessment programme was designed not only to report on the achievement levels of Thai students but to encourage teachers and administrators to widen the scope of their lessons beyond the needs of students who would sit for university entrance examinations. Like the new curriculum itself, the assessment programme emphasized affective learning outcomes such as attitudes toward work, moral values and social participation. The assessment has been credited with helping teachers in Thailand to realize the importance of affective learning outcomes.[51]

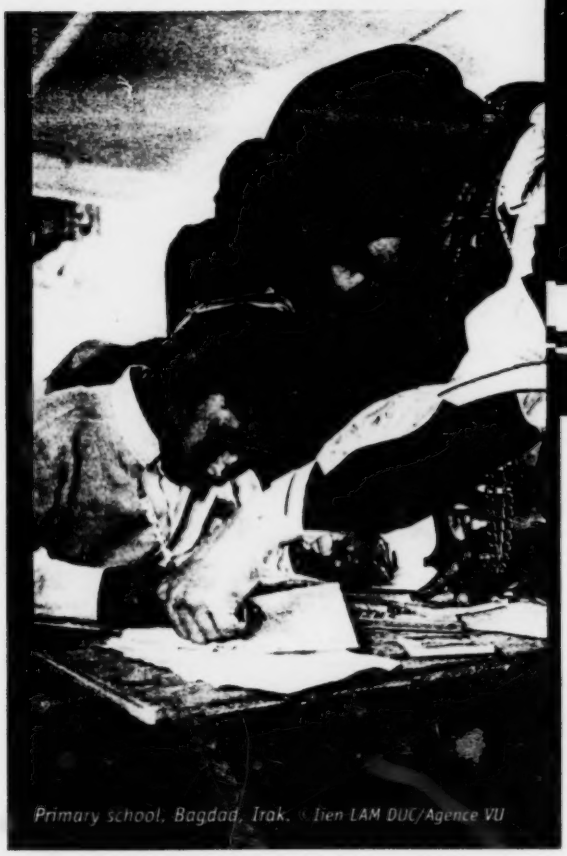
Various countries have used assessments to affect teaching. Lesotho constructed skills checklists and booklets with sample test questions to accompany a new core curriculum, while Chile looked at the affective domain. Like Thailand, China has used its assessments to put more emphasis on the academic needs of students not going on to universities. China followed this course after discovering that while pupils were doing reasonably well in learning to read, write and do mathematics, their knowledge of life skills was

49. Greaney and Kellaghan, 1996, op. cit., p. 18.

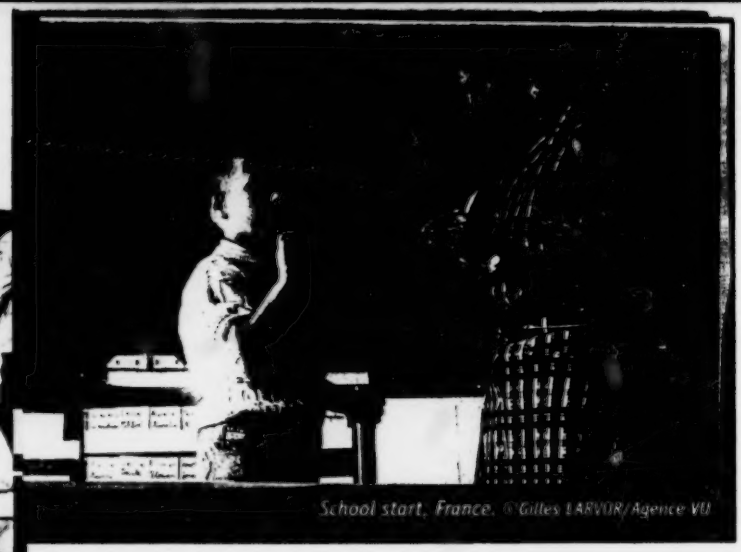
50. Wolff, op. cit., p. 6.

51. Pravalpruk, Kowit, '*National Assessment in Thailand*,' in *National Assessments: Testing the System*, edited by Paud Murphy, Vincent Greaney, Marlaine E. Lockheed, Carlos Rojas, Washington DC, The World Bank, Wash., DC, 1996.

BEST COPY AVAILABLE



Primary school, Bagdad, Irak. © Tien LAM DUC/Agence VU



School start, France. © Gilles LARVOIR/Agence VU

lagging. Zanzibar used data from its MLA research to promote active learning strategies and a child-centered approach to teaching.

In some cases, the impact of assessments can be measured in negative terms. In Kenya, for example, officials found that subjects such as health, nutrition and agriculture, which were not tested, received little attention in classrooms.[52] Teachers in Ethiopia, Madagascar and Zambia have been reported to have ignored the teaching of practical subjects because such topics are not tested.

Influencing textbooks and curricula

Assessments can provide valuable feedback to authors of textbooks. Designers of instructional materials in Chile used SIMCE results to identify areas in which students were having problems and to address these areas in preparing new and revised materials.

In addition to affecting classroom teaching, the assessment programme in Thailand had a major impact on curriculum development. Results of the assessment snowed that students were performing at an unsatisfactory level in both science and mathematics, especially when it came to applying this knowledge to practical problem. Scores were also found to be declining in other subjects in the area of thinking skills. The Department of Curriculum and Instructional Development used these quantitative data in 1989 to revise the curriculum in order to emphasize process skills.[53]

Kenya altered the content of public examinations at the end of primary schooling to include a much broader spectrum of cognitive skills than in the past and to promote the teaching and learning of competencies useful to the majority of pupils who were likely to leave school after the examination. Emphasis was simultaneously shifted from measurement of pupils' ability to memorize factual information to the testing of comprehension and application of knowledge.[54]

Building public support for education

As already noted, Chile used the results of the SIMCE assessments to bolster public awareness of education. Student performance was reported by school, location and region, and results were publicized through newspapers, radio, television, posters and videos. Manuals were published to explain the results, and parents used test scores when deciding where to send their children to school. As a result of these efforts, the assessment programme enjoys widespread public support.

52. Heyneman and Ransom, op. cit., p. 186.

53. Pravalpruk, op. cit., p. 145.

54. Greaney and Kellaghan, 1992, op. cit., p. 13.

Other countries that have successfully used assessments to build support for public education include Kenya and Sudan, where the results of national public examinations are published annually in order to provide parents with an opportunity to compare schools.[55] On the other hand, Mexico has demonstrated a longstanding reluctance to publicize examination results.[56]

One danger of publishing results is that parents and others may use them to judge schools unfairly.

Assessing learning readiness

Costa Rica has given initial diagnostic assessments to 10 per cent of children entering first grade in order to develop information on their physical, cognitive and social-emotional status. These data are then to be used to establish guidelines for appropriate learning experiences in pre-school and first grade classrooms.[57] ■

TESTING FOR THE BENEFIT OF LEARNING IN TUNISIA

In 1995/96, the Tunisian Ministry of Education, in cooperation with UNICEF, launched the experimental phase of the "Basic Competencies Programme" which was aimed at improving the effectiveness, equity and quality of the educational system.

The BC Programme sets up hierarchies of learning objectives, ranging from entry level competence to desired outcome competencies. Teaching is structured to take pupils from one level of competence to another through a carefully defined approach which prepares the child to mobilize and integrate his various competencies in situations of communication or of problem-solving. The competency-based teaching programme proceeds therefore from a holistic view of learning which places the child at the centre. Under this approach, learning is participatory and the traditional roles of teacher and pupil are significantly changed.

Systematic monitoring of child performance and progress is a crucial dimension in the programme which adopts the approach of "formative evaluation". The monitoring system is based on a testing procedure with 4 tests a year taken by the students of each grade in Math, Arabic and French. The test is criterion-referenced and assesses the competencies acquired by the students during the year. This regular performance assessment allows the teacher to detect the specific difficulties faced by children and to analyze the source or origin of the error. The student will work then on a series of well-targeted exercises to overcome the difficulties encountered: the previous competencies are reviewed and consolidated before moving on to new ones. The remedial work may be undertaken in small groups or through child to child tutoring methods.

This testing procedure has become the key instrument for teachers to orient, regulate and evaluate their work and for pedagogic supervisors to detect fields where corrective measures need to be taken.

Since the launching of the project, a computerized data base has been set up with a sample of 50 experimental schools and a group of 20 control schools. The students' results in the 4 tests are entered in the data base and then processed so as to review their learning achievements. A report presenting the overall progress of students' performance in the BC schools since 1995 will be published this year.

An external evaluation of the Programme was jointly carried out by the MOE and UNICEF in April 1999. The results of the evaluation and of the data base analysis were presented in a meeting led by the Minister of Education with the participation of all national and regional directors, advisors, UNICEF and the World Bank. This meeting concluded with the decision to progressively generalize the competency based approach to national scale both in primary and in early secondary education.

Source: UNICEF, Tunisia.

55. Greaney and Kellaghan, op. cit., p. 178.

56. Wolff, op. cit., pp. 10-11.

57. Wolff, op. cit., p. 7.

PART VII

Issues of measurement

In setting out to use assessments as a means of improving student achievement, policy-makers must deal with a number of practical issues. Following is a discussion of some of these issues.

What to assess

The most fundamental questions facing policy-makers as they set out to design a national assessment concern the academic materials to be covered. Most countries focus assessments on reading, writing and mathematics and, possibly, other core subjects such as science. Some ministries of education, however, are also interested in how students perform in affective domains. Colombia measured attitudes toward school, while the SIMCE programme in Chile evaluated pupils' self-esteem.

In recent years a number of countries have attempted to evaluate life skills, (see box p. 47) and the MLA project put considerable emphasis on developing techniques for doing so. China, for example, included a number of questions aimed at determining how well children were prepared to deal with accidents (such as oil catching fire in a kitchen) and how much autonomy they showed on matters such as changing light bulbs and choosing the right tools.[58]

Even when policy-makers have decided on the particular domains to measure, there are important decisions to be made. As seen in the previous sections, officials in China, France, Thailand and other nations have made conscious decisions to broaden the range of questions so as to focus on the needs of the vast majority of students, not merely those who are pursuing a narrow academic track in order to qualify for universities. Such decisions reflect both practical and equitable considerations.

Another set of decisions with profound equity consequences revolves around whether to base assessments on the official curriculum or on what is actually covered in classrooms. The former approach will benefit students who attend schools where teachers are able to get through all or most of the prescribed topics, but such an approach could transform the assessment into a measure of students' opportunities to learn rather than a measure of actual learning.[59]

Sampling

National assessments are undertaken for the purpose of estimating achievement levels for the system as a whole rather than for individual students. While it would seem logical to assess an entire cohort of pupils in particular subjects – such as all Grade 4 pupils in reading and mathematics – such an approach is usually neither feasible nor necessary. Most countries achieve their purposes by assessing a sample of students. Colombia, for example, has developed a "master sample" that will be maintained

58. Chinapah, op. cit., pp. 112-15.

59. Heyneman and Ransom, op. cit., p. 186.

63

BEST COPY AVAILABLE

REQUIREMENTS OF A GOOD SURVEY

Poorly-designed and poorly-executed studies provide little that is of use to Ministries because they cannot be used in any meaningful way for policy analysis and development*

In view of the growing number of international studies of educational achievement, the International Academy of Education was asked to prepare a document for Ministries of Education that would not only assess the benefits and limitations of such studies, but also alert readers to the kind of questions they should be asking themselves when reading the reports of such studies.

To conceptualize, conduct and interpret international surveys is far from being an easy matter and Beaton and his collegues list a number of concerns, related mainly to quality and relevance. They suggest that a good study should stand the test of the following questions:

- Aims of the study: are the research questions and results really responding to them?
- Design of the study: was the sampling conducted to result in accurate estimates?
- Target population: what about the 'excluded' population, such as pupils living in very isolated areas? Is it less than 5 per centof all pupils in the 'desired' target population, as usually required?
- Sampling: are the standard errors of sampling given in the tables of results? Are they acceptable for the purposes of the study?
- Instrument construction: how do the test instruments cover different national curricula? What about problems of bias and translation?
- Data collection: are the levels of missing schools less than 10 per cent and the level of missing students less than 20 cent? These are the absolute maximum levels.

As to what meaningful comparisons can be made within and between countries, one always has to remember the complex and multi-factorial nature of the education process. The rationale is to ask, others things being equal, what is the impact of this particular factor on education achievment? Are differences within and between countries and/or changes over time due to chance or not, and if not, how important are they? A large number of students in surveys are apt to render any difference statistically significant, which does not mean that it is meaningful! These differences can also be due to variations in education systems or in assessment methodologies.

Factors can be invariant within countries but vary among countries, such as the starting age of formal schooling. When comparing subgroups as girls and boys, comparisons should include not only achievement levels but such factors as enrolment rates and attitudes towards learning (see box p. 42). Sometimes, the potential influence of certain factors such as teaching methods and attitudes can better be examined in replicated small experiences than aggregated national data for a few countries. And last but not least, results have to be put in context according to the educational and cultural context of the country. Comparisons have to be fair, they have to 'assure the measure and respect the contexts'**.

* Beaton et al. *The Benefits and Limitations of International Educational Achievement Studies*. 1999. IIEP/UNESCO. Paris.

** P. Vrignaud, and D. Bonora. *Evaluation de la Litteratie et Comparaisons Internationales: Assurer la Mesure, respecter les Contextes*. Séminaire de l'UNESCO. 22-24 juin. 1998. Paris.

over a ten-year period using 5,000 schools out of a universe of 50,000.[60] The MLA project in Nigeria surveyed 23,040 pupils in 960 schools using a stratified random sample of 24 pupils from each of 32 schools in each of the 30 states of the Federation as well as 24 pupils from each of 16 schools in the Federal Capital Territory of Abuja.[61]

Use of sampling cuts overall costs, permits greater speed in retrieving and analysing data, and makes it possible to have more in-depth supervision of those carrying out the assessment. Defining and measuring an accurate and systematic sample of pupils, however, requires careful planning. Typical obstacles include the absence of a centralized list of pupils in a

60. Margarita Pena. *The National Assessment System in Colombia*.

61. Falaxaio et Al. op. cit.

BEST COPY AVAILABLE

particular cohort and even the existence of 'phantom' schools. In many countries officials are reluctant to use sampling because of fears that it will undermine public confidence in the accuracy of the assessment. Chile has resisted using sampling and the resulting increase in costs was seen as one reason why the assessment programme was abandoned temporarily in the mid-1980's.[62] On the other hand, when sampling is employed, teachers whose classes are selected for testing may conclude that they are being singled out for disciplinary or other reasons.

Even when such technical and political problems are solved, examination officials must deal with complex issues such as whether to exclude certain types of students, such as those with learning or language difficulties or even entire schools, notably those in remote regions of the country. The use of matrix sampling, whereby each pupil completes only a fraction of the test, offers additional benefits but also poses its own particular technical problems.[63] Educators organizing assessments are usually well advised to seek the co-operation of census officials in order to benefit from their expertise.

Quality of data

Assessment programmes in all countries must deal with qualitative issues of reliability, validity and the comparability of data over time. Such issues are particularly important in developing countries where resources are scarce and where it is important to get the maximum return from those that are invested in assessment. It is important that data be reported with the standard errors of sampling.

Quality problems have arisen in many of the recent assessment activities. In Chile, for example, authorities have not taken steps to ensure that the SIMCE test results are compatible from year to year. This omission makes it difficult to compare the performance of schools over a period of time. Likewise the compatibility of the previous MLA studies has been reduced because each country was allowed to modify the design. The major international comparative studies have not been immune to such problems. In releasing their findings, TIMSS researchers put asterisks next to the names of countries that did not comply with all of the procedural standards. The quality of assessment data also varies among the various countries that participate in multi-national efforts such as the Laboratorio, MLA and SACMEQ.

The SACMEQ project has shown that research of world class quality can be carried out in developing countries. The standards of its research procedures – including design, sampling, test and questionnaire construction, trial testing and field operations – have exceeded that of most previous large-scale educational survey research in Africa, including that conducted by measurement experts from developed countries. In fact, some of the technical training materials produced by SACMEQ, such as its sample selection software and training manuals for sample design, have been adopted by organizations such as IEA.

Security

Assuring the integrity of examinations is central both to obtaining reliable data and to maintaining public confidence in the assessment process. Developed and developing countries alike have been forced to deal with acts of cheating ranging from smuggling notes into examination rooms and passing information during the course of examinations to the illicit purchase of entire examinations and the physical intimidation of test monitors and scorers. In April 1999, for example, the deputy superintendent of the Austin, Texas, school district in the United States was indicted for allegedly tampering with test data of low-performing students in order to raise scores on state examinations.[64]

Many students are understandably tempted to cheat on high stakes examinations where the score they obtain will determine whether they continue in school or leave for a low-paid jobs, and teachers and school administrators face their own temptations. The professional reputation of teachers and funding levels of schools often hinges on how well their pupils do on important examinations. In Chile, for example, there was some evidence that schools overestimated the levels of poverty among their students in order to enhance their standing on the SIMCE assessments.[65]

Research suggests that the level of malpractice varies greatly from country to country. Some of the variation reflects cultural differences; what is dishonest in one context may be viewed as helping a friend in another. But levels of malpractice are also a

62. Greaney and Kellaghan, 1996, op. cit., p. 18.

63. Ibid, pp. 47-50.

64. *Education Week*, April 14, 1999.

65. Greaney and Kellaghan 1996 p. 19.

THE IMPORTANCE OF LANGUAGE POLICIES: PAPUA NEW GUINEA AND PARAGUAY

Papua New Guinea, a nation with about 4,5 million people, has 820 languages, of which 3 are national languages or languages of wider communication (English, Tokpisin and Hiri Motu), while 17 have fewer than 50 speakers each*. In 1995, this small island nation amended its Education Act to declare that all languages may be used for elementary education, the first three years of formal education.

This innovative programme grew out of a successful non-formal, community based and community driven non-formal vernacular programmme. In his research, Jonduo** found that children achieved better and stayed in school longer when they started their education in a language they spoke.

Since the launching of the first elementary classes in Milne Bay Province in 7 languages in 9 schools, another approximately 247 languages have started elementary programs in all 20 of the nation's provinces. The first of the children concerned are now in Grade 6. In the year 2000 it is estimated that 190,000 children are enrolled in either elementary prep, one or two.

Anecdotal evidence and some statistical data are emerging which show a significantly improved level of achievement for these children. Patricia Paraide, with the Department of Education's Facilitation and Monitoring Unit, says, "Children in elementary schools take an interest in their learning. They are able to take an active role in their learning process."

The challenges are huge. Half the nation's languages do not have written alphabets. The Department of Education, in co-operation with SIL PNG and other organizations in the country, is working to develop alphabets. The Department of Education supports curriculum development by providing teacher resource materials and student materials which local communities can adapt to their own languages.

Many educational leaders in Papua New Guinea feel that when people lose their language, they lose their culture. Then they lose their identity. Papua New Guinea believes that it is worth the effort and the cost to start children's education in a language they already speak, and then add skills in a national language.

In Paraguay, by contrast, 87 per cent of the population speak Guarani***, the common everyday language at home, in the office, in the market and even of electoral campaigns. While half of the population uses both languages, Spanish, was until recently the only language of the education system. But in 1992, as the country opened to democracy, this Amero-indian language became recognized as equal to Spanish and a totally bilingual education system was established.

The bilingual education plan was launched in 1994, starting with first grade. Children learn how to read in their own language but the second language is introduced from the beginning with a gradual progression during the nine years of compulsory schooling. Both Guaraphone and Spanish-speaking children share the same curriculum. The objective is to develop truly bilingual speakers "bilingües coordinados" with equal skills and knowledge in both languages.

The programme, which is national, compulsory and non-religious, is conceived in such a way as to comprise all levels of the education system, including university. Priorities include teacher training, preparation of textbooks in Guarani and Spanish, as well as furnishing school libraries with books in both languages. Exam equivalencies are sought with other Mercosur countries.

The first evaluation reports**** suggest that the usage of Guarani improves communication in school as well as the self-image of children, and that bilingual pupils demonstrate more agile and lively minds. Parents are also more involved in schools and recent studies show that learning achievement has increased by 15 per cent since the system was introduced.

* Landweer, M. Papua New Guinea Language Status, unpublished manuscript, 1999.

** Jonduo, W. Effectiveness of Tokples Skul Tisa Training. Policy, Communication and Research Branch. National Department of Education. Papua New Guinea,1993.

*** Delicia Villagra-Batoux. Le guarani paraguayen : de l'oralité à la langue litteraire, Septentrion. Presses universitaires de Lille. France, 2000.

**** Ministerio de educacion y cultura. Evaluation en el ano 2000. Informe de Paraguay, version preliminar, setiembre de 1999. Asuncion.

BEST COPY AVAILABLE

reflection of conscious policy. Examination officials can take obvious steps to assure the security of testing sites, such as guarding copies of the examinations beforehand and providing adequate monitors, but other protective steps are more complicated. Cheating may be more of a temptation in situations where students perceive a disconnect between hard work and success on the examinations. Students who believe that their schooling has left them poorly prepared for a test may see no alternative to malpractice. Likewise, teachers confronted with overly ambitious syllabuses or insufficient resources may be tempted to look for ways to give their students an edge.[66]

Political influences

Although a strong case can be made on educational grounds for using assessments to improve teaching and learning, political forces sometimes intervene to limit such use. In some cases, such as those where student performance is low, political considerations become reasons to block publication of results. In Mexico, where there is a long tradition of reluctance to publish examination results, officials declined at the last moment to release the country's results on TIMSS.[67] Countries that view themselves as competing with particular neighbors, such as those on the Indian sub-continent, are understandably reluctant to place themselves in a position where their students might be seen as less able than those in the rival nation.

Political forces can also become important to the extent that assessment activities succeed in encouraging educational reform. As Vincent Greaney and Thomas Kellaghan observed, "Any proposed reform must confront the potentially conflicting interests and values of the different stakeholders in the education system." These stakeholders range from religious and ethnic groups to teachers unions, school administrators and textbook publishers. At the same time, they add, 'National assessment data on the fuctioning of a school system can plan an important role in ensuring that the selection of educational priorities does not depend solely on the values of politically powerful groups.'[68]

Lack of understanding of the potential impact of assessment on educational quality can also be a problem. In Colombia, for example, managers of National Assessment System (NAS) made great strides in building a technically reliable system, and during its early stages considerable attention was paid to the dissemination of results both to government policy-makers and to the public as a whole. In recent years, however, both the frequency and the quality of dissemination efforts have declined, in part because, as one observer put it, top government officials have only a 'limited understanding ... of the potential of this information as a tool for quality improvement.' Another factor, she suggested, may be that, since the assessment has been shaped by measurement specialists rather than policy-makers, NAS is 'a system more interested in technical developments than in educational improvement.'[69] ■

66. Greaney and Kellaghan (Ch. 12), p. 177-9.

67. Wolff, pp. 10-11.

68. Greaney and Kellaghan 1996 p. 67.

69. Pena.

67

Conclusions

The experience of developed and developing countries alike in carrying out national assessments has led to a body of general principles about how best to organize such measurement exercises and how to make the most effective use of the data they generate. Here are some of these principles:

Clarity of purpose

Assessments of student learning can be organized in a multitude of ways and serve a variety of purposes, and no single approach will accomplish all objectives. In mounting an assessment project, national policy-makers must clarify the goals they hope to reach and then design a measurement approach appropriate to these goals. The process of organizing a national assessment can, in turn, help policy-makers develop a better sense of curricular, performance and other goals. For example, if the most important goal is to map overall achievement at the national level, then an assessment system based on systematic sampling of pupils at key points in their educational process is most likely warranted. If the goal is to assess strengths and weaknesses of particular pupils and to provide feedback to classroom teachers, then a quite different model is required. As one observer wrote, 'Failure to design and implement an assessment program that is consistent with national purposes will result in wasted resources and in the inability to make policy, resource allocation, and instructional decisions that will have the most educational impact.'[70]

In some situations legitimate goals may be contradictory. It is one thing for policy-makers to use assessments to measure how well various reform measures are working. It is quite another to use them to drive the reform process in the first place.

In any case, the very act of examining the purposes for mounting assessments can be a constructive activity

Underlying values

The design of any assessment programmes involves assumptions – stated or unstated – about the fundamental values that underlie any state education system. If the assessment is to accomplish its goals, it is important that these assumptions are transparent and that there be general agreement among various stakeholders that they are fair and reasonable.

One of the most important value judgements typically involves the target population of test takers. Is the assessment designed to enhance the education of pupils or students pursuing core academic subjects with the aim of continuing their education at higher levels, or is it designed to serve the learning needs of the broader student population? Such issues are equally important in the designing of public examinations, which can serve either the limited purpose of selecting an elite group of pupils to continue their education or the broader purpose of certifying larger numbers of pupils for jobs. Closely related are questions of whether to test out-of-school as well as in-school knowledge and experience and whether to assess higher order thinking skills.

Other value judgements revolve around equity issues, such as whether the resulting assessment data will be organized in such a way as to highlight learning problems in particular regions or among particular groups of pupils.

70. Williams, Paul L. *Trends and Issues in Large Scale Student Assessment Programmes.* Education Update. UNICEF. January 1999

68

Linking assessment to improved teaching and learning

The ultimate purpose of any national assessment exercise must be to improve the quality of teaching and learning. This 'bottom line' should be kept in mind in designing each aspect of the measurement process.

A focus on quality improvement requires designers to understand the extent to which assessments can drive classroom teaching. Assessments must continually evolve in order to keep up with changes in curricula, textbooks and teacher preparation. Well thought-out feedback mechanisms directed at pupils, teachers, school administrators and the community as a whole, are necessary.

Policy-makers and managers must also be alert to negative effects that assessment can have on teaching and learning, such as narrowing the range of topics taught by classroom teachers.

The usefulness of assessments in improving pupil/student achievement depends to a large extent on the kind of information that is reported. Assessments that probe topics such as student backgrounds and teacher attitudes and preparation are more useful than those limited to measuring student knowledge alone. Likewise, assessments that permit analysis of critical thinking or problem-solving skills are more useful to teachers and administrators than those that supply more generalized information.

Capacity-building and technical competence

In the long run the success of any country's assessment programme will depend on whether that nation is able to build up its own internal technical capacity to design and carry out assessments suited to its particular needs. While many of the techniques for effective measurement are applicable to all countries, each nation poses its own particular challenges because of differing cultural, political and social climates. In the absence of local professional expertise, assessments will never fulfil their potential.

Participating in cross-national assessment projects – with the resultant infusion of knowledge, training and experience – can be an effective and cost-efficient way to jump start this process, but it is only a first step. Assembling an inter-disciplinary team of professionals takes time. A long-term commitment to capacity-building requires quality training of assessment experts, possibly through new degree programmes in measurement, and it requires the development of an ongoing research enterprise aimed at updating and renewing the programme. Likewise, training classroom teachers to make effective use of continuous assessment is a long-term enterprise.

The MLA project offers an effective model for spreading expertise throughout countries. Central officials organize a limited number workshops for assessment workers, who in turn fan out and train their colleagues. In China, sixty workshops led to the ultimate involvement of 6,500 persons.

Heyneman and Ransom suggest that national testing agencies should be judged both by the quality of the tests they develop and the nature of the feedback that these tests produce. In addition, 'an effective testing agency must be able to make professional decisions on politically and professionally controversial and sensitive issues, such as the language of assessment, test coverage and fairness, and appropriate testing for practical subjects.'[71]

The limitations of assessment

It is important for managers of national assessment programmes to understand the limitations as well as the potential of such programmes. The most obvious limitations stem from structural problems in the school system itself. The usefulness of such programmes can be limited by a shortage of places in secondary schools, poor quality teaching, inadequate supplies, language problems and other forces beyond the control of assessment authorities. Assessment is not a stand-alone enterprise. It must be developed in tandem with improvements in curriculum, teacher training and the other elements of a state education system. Likewise, neither educators nor political leaders are likely to pay much attention to the results of the assessment if they do not have confidence that it is a quality enterprise.

Assessments can have negative as well as positive consequence. Critics of ministry of education policies, for example, can seize on low scores as a weapon to serve their own political purposes

71. Heyneman and Ransom, op. cit., p.191

Examinations can have the effect of limiting the topics that classroom teachers cover in their instruction, and test questions that stress recall and recognition of factual information rather than broader thinking skills will send the wrong sort of messages to teachers.

Financing assessments

Developing countries must make difficult choices about how to allocate scarce resources. In nations whose state education system lacks adequate school buildings, textbooks and other basic resources, it is reasonable to ask whether it is appropriate to divert funds into assessment activities.

On the other hand, as Greaney and Kellaghan point out, it can be argued that the resources required to carry out a national assessment would not go very far in solving such problems and the information obtained from assessment can be used to make the system more cost-effective. 'It is up to the proponents of a national assessment to show that the likely benefits to the education system as a whole merit the allocation of the necessary funds' they write. 'If they cannot show this, the resources earmarked for this activity might be more usefully devoted to activities such as school and textbook provision.'

Dissemination and public support

If assessments are to enhance student achievement, there must be a substantial degree of national consensus about their value. There must be strong political will among national leaders at the outset and there must be a commitment to long-term support of a process that may bring disturbing as well as positive news, especially at the outset. In the absence of political will and long-term commitment, initial funding for measurement is likely to dissipate and be reallocated to other educational priorities. This is what happened to the PER programme that preceded the SIMCE programme in Chile.

Regular and timely public dissemination of assessment results must be built into the process from the very outset and dissemination must take forms that are readily understood by the public at large, especially the release of statistics. Chile had to abandon the use of percentiles, for example, because their meaning was not readily grasped.

Experience shows that efforts to involve the public in determining standards are helpful. In any case, continued public confidence in the quality and fairness of tests is essential to any assessment programme. As Mehrens writes, "The public and the press are more likely to use what they believe to be 'inadequate' assessment results to blame educators than to use 'good' results to praise them."[72]

Future prospects

The insights described above regarding the assessment of learning achievement within the context of formal education can be extended to broader concerns about the levels of literacy, life skills, and social and employment skills of the general population in the rapidly changing global society.

Frequent calls are now being heard to shift the basis of such assessments from a curriculum-oriented to a competency-based approach. Instead of assessing student performance in relation to defined school curriculum objectives, the focus could be on the extent to which youth and adults have acquired the fundamental generic competencies that are necessary for them to function in, and benefit from, modern society. Assessment of these competencies, which may be acquired either in or out of school, would begin with basic literacy and numeracy. It would also cover capacity for creativity and the ability to search for, access and understand information; to make sound judgements about the relevance and truth of the information; to use the knowledge obtained in solving problems; and to communicate information and ideas.

Building on recent initiatives under the International Adult Literacy Survey (IALS) to assess functional literacy among the adult population in a number of industrialized countries, the National Center for Education Statistics (NCES) in the United States has launched a new initiative to promote the assessment of basic competencies. The initiative, called the International Life Skills Survey (ILSS), represents the opening of a new page in the assessment of individual learning outcomes. When dovetailed with the on-going assessment programmes described above, this initiative promises to contribute not only to a better understanding of the learning phenomenon but also future strategies for building a truly learning society. ■

72. William A. Mehrens, Consequences of Assessment: What is the Evidence? Education Policy Analysis Archives, Vol. 6 No. 13, 4 1998.

Annexes

Introduction to the statistical tables

The tables in this appendix are based largely on data reported to the UNESCO Institute for Statistics and to other international agencies. While these data are the best available at the international level, they must be interpreted with care and should be regarded as indications of magnitude rather than as precise measurements. At the international level, pertinent and comparable data on learning achievement are readily available only for a very few number of countries. Most of these data are presented and discussed inside this report. As for the appendix, the selected data describe aspects of the education systems (in the first five columns) and certain social and economic indicators (in the remaining columns to the right) that may have an effect on conditions of teaching and learning in different countries. Due to important differences between countries, especially with regards to their school, social, cultural and economic systems, one should be prudent when comparing data across countries or attempting to rank countries by any single indicator.

The Net Enrolment Ratio (NER) (column 1) is the number of pupils in the official primary school-age group expressed as a percentage of the total population in that age group. NER is calculated for the entire primary education cycle, using the country official school-age group which varies in length from country to country. *The Percentage of first graders reaching Grade 5* (column 2) includes all children having started school in 1995 and who are expected to reach Grade 5 with or without repetition. *The Percentage of Repeaters* (column 3) is calculated for the entire primary education cycle. *The Estimated Adult Literacy Rates* (column 4) refer to the population aged 15 and over. *The Public Current Expenditure on Pre-primary and Primary Education* (column 5) refers to percentages based on government (central, regional, local) current expenditure, i.e. excluding capital expenditure, in relation first to total current expenditure on all levels of education, then to gross national product (GNP) per capita. These two indicators give a rough measure of a country's financial effort in favour of the major components of basic education in relation to its means.

Column 6 shows estimates of *GNP per capita in US dollars*. Column 7 shows an indicator of poverty – the *Percentage of people living on less than one dollar per day*, which is a calculated average over a 15-year period (1981-1995) expressed in 1995 international dollars, i.e. based on purchasing power parities (PPP). Using PPP-based currency values in international comparisons has the advantage of overcoming the problem of using market exchange rates, which do not always reflect a currency's true purchasing power at home. A PPP-based currency depicts how much it can purchase locally of a common 'market basket' of goods and services, including goods and services that are not traded internationally. The *Estimated percentage of economically active children, 10-14 years old* (column 8), gives a rough measure of an important cause of non-enrolment, school drop-out and pupil low achievement. The *Daily newspaper per 1,000 inhabitants* (column 9) and *Television receivers per 1,000 inhabitants* (column 10) show the level of communication and cultural development of a society. Column 11, *Percentage of urban population*, refers to national criteria of urban/rural area. In general, urbanisation is associated with more school facilities and better conditions of teaching and learning. *The Total fertility rate* (column 12) is the number of live births in a given year per 1,000 women in the group-age 15-49 years. This indicator tend to decline as the level of education of women rises. *The Life expectancy at birth* (column 13) is the average number of years in new-born infant (in 1997) is expected to live if prevailing patterns of mortality in the country remain the same throughout the individual's life. ■

SUB SAHARAN AFRICA

COUNTRIES/ TERRITORIES	NET ENROLMENT RATIO				PERCENTAGE OF 1995 COHORT REACHING GRADE 5		PERCENTAGE OF REPEATERS 1996		ESTIMATED ADULT LITERACY RATES 1997				PUBLIC CURRENT EXPENDITURES IN PRE-PRIMARY AND PRIMARY EDUCATION[1]		GNP PER CAPITA US $ 1997	% PEOPLE LIVING ON LESS THAN $1 A DAY (PPP) 1981-1995	ESTIMATED % OF ECONOMICALLY ACTIVE CHILDREN (10-14 YEARS)	DAILY NEWSPAPER PER 1,000 INHABITANTS 1996	TELEVISION RECEIVERS 1,000 INHABITANTS 1997	% URBAN POPULATION 1996	TOTAL FERTILITY RATE	LIFE EXPECTANCY AT BIRTH	COUNTRIES/ TERRITORIES
	1990		1996						15-24 YEARS		15 YEARS AND OVER		AS % OF TOTAL CURRENT EXPENDITURE ON EDUCATION	PER PUPIL AS % OF GNP PER CAPITA							1997	1997	
	BOYS	GIRLS	BOYS	GIRLS	BOYS	GIRLS	BOYS	GIRLS	MALE	FEMALE	MALE	FEMALE											
ANGOLA															260		27	11	13	31	7	46	ANGOLA
BENIN	65	32	80	47	64	57	25	25	67	33	48	21	59	12	380		27	2	11	38	6	53	BENIN
BOTSWANA	90	97	79	83	87	93	4	3	83	91	72	77			3 310	35	16	27	20	60	4	47	BOTSWANA
BURKINA FASO	33	21	37	24	74	77	16	16	41	18	30	11	57	20	250		48	1	9	16	7	44	**BURKINA FASO**
BURUNDI									64	57	54	36	43	20	140		49	3	4	8	6	42	BURUNDI
CAMEROON									92	90	79	65	**87**	**11**	620		24	7	32	45	5	55	CAMEROON
CAPE VERDE									90	81	82	62			1 090		14		4	54	4	69	CAPE VERDE
CENTRAL AFRICAN REPUBLIC	64	42							73	52	56	30	53	8	320		30	2	5	39	5	45	**CENTRAL AFRICAN REPUBLIC**
CHAD			59	33	62	53	32	32			64	37	44	6	230		38	0	1	22	6	47	CHAD
COMOROS									65	52	63	48	37	9	400		39		2	30	5	59	COMOROS
CONGO					40	78	35	32	97	95	85	70	50	15	670		26	8	12	58	6	49	CONGO
COTE D'IVOIRE			63	47	77	71	24	25	67	54	51	34	45	17	710	18	20	17	64	43	5	47	COTE D'IVOIRE
DEM. REP CONGO	61	48													110		29	3	135	29	6	51	DEM. REP CONGO
EQUATORIAL GUINEA									98	94	91	70			1 060		33	5	10	42	6	50	**EQUATORIAL GUINEA**
ERITREA			32	29	73	67	19	22							230		39		0	17	6	51	ERITREA
ETHIOPIA			39	24	51	50	7	9	53	49	41	29	46	25	110	34	42	2	6	15	6	43	ETHIOPIA
GABON					58	61	36	34			76	57			4 120		17	30	55	50	5	52	GABON
GAMBIA	60	42	72	57	78	83	13	12	61	44	40	26	49	13	340		36	2	4	29	5	47	GAMBIA
GHANA									91	82	77	56			390		13	14	93	36	5	60	GHANA
GUINEA			50	33			27	29			52	24	35	8	550	26	33		12	29	6	47	GUINEA
GUINEA-BISSAU									68	34	50	18			230	87	38	5		22	6	45	GUINEA-BISSAU
KENYA									95	92	87	72			340	50	40	9	26	29	4	52	KENYA
LESOTHO	65	81	64	76	55	71	23	18	82	98	71	93	41	14	680	50	22	8	27	24	5	56	LESOTHO
LIBERIA									82	49	65	32					17	16	29	45	6	47	LIBERIA
MADAGASCAR			59	62	49	33	35	33					30	4	250	72	35	5	22	26	5	58	MADAGASCAR
MALAWI	52	48	100	100	36	32	15	15	80	58	73	43	59	9	210		34	3		13	7	39	MALAWI
MALI	27	16	38	25	92	70	18	18	66	51	43	28	46	14	260		53	1	4	27	7	53	MALI
MAURITIUS	95	95	98	98	98	99	5	4	92	93	87	79	31	8	3 870		3	76	228	41	2	71	MAURITIUS
MOZAMBIQUE	49	38	45	34	52	39	25	27	72	41	57	25			140		33	3	5	34	6	45	MOZAMBIQUE
NAMIBIA							14	12	89	92	81	79	58	21	2 110		20	19	37	36	5	52	NAMIBIA
NIGER	32	18	30	19	72	73	13	13	29	12	22	7			200	62	45	0	13	18	7	49	NIGER
NIGERIA									87	80	69	51			280	29	25	27	66	40	5	50	NIGERIA
RWANDA	66	66							83	78	71	56			210	46	42	0	0	6	6	41	RWANDA
SAO TOME ET PRINCIPE															290				163	43			SAO TOME ET PRINCIPE
SENEGAL	55	41	65	55	89	85	13	13	56	38	45	25	34	11	540	54	30	5	41	44	6	52	SENEGAL
SEYCHELLES					98	99							27	10	6 910			46	145	54			SEYCHELLES
SIERRA LEONE											47	20			160		15	5	12	33	6	37	SIERRA LEONE
SOUTH AFRICA	100	100	95	96					90	90	85	83	44	15	3 210	24	0	34	134	49	3	55	SOUTH AFRICA
SWAZILAND	87	88	90	91	73	79	18	13	88	90	79	76	36	7	1 520		13	27	23	31	5	60	SWAZILAND
TOGO	87	62	93	69	79	60	24	25	85	59	69	38	46	9	340		28	4	17	31	6	49	TOGO
UGANDA									84	69	75	53			330	50	45	2	16	13	7	40	UGANDA
UNITED REP. OF TANZANIA	51	52	48	49	78	84	2	2	92	85	82	62			210	16	38	2	3	24	5	48	UNITED REP. OF TANZANIA
ZAMBIA			76	74			3	3	90	82	83	67	42	5	370	85	16	14	32	43	6	40	ZAMBIA
ZIMBABWE					78	79			100	98	94	88	52	18	720	41	28	19	33	32	4	44	ZIMBABWE

1. Bold figures include secondary education

BEST COPY AVAILABLE

LATIN AMERICA/CARIBBEAN

COUNTRIES/ TERRITORIES	NET ENROLMENT RATIO				PERCENTAGE OF 1995 COHORT REACHING GRADE 5		PERCENTAGE OF REPEATERS 1996		ESTIMATED ADULT LITERACY RATES 1997				PUBLIC CURRENT EXPENDITURES IN PRE-PRIMARY AND PRIMARY EDUCATION*		GNP PER CAPITA US $ 1997	% PEOPLE LIVING ON LESS THAN $1 A DAY (PPP) 1981–1995	ESTIMATED % OF ECONOMICALLY ACTIVE CHILDREN (10–14 YEARS)	DAILY NEWSPAPER PER 1,000 INHABITANTS 1996	TELEVISION RECEIVERS 1,000 INHABITANTS 1997	% URBAN POPULATION 1996	TOTAL FERTILITY RATE	LIFE EXPECTANCY AT BIRTH	COUNTRIES/ TERRITORIES
	1980		1996						15–24 YEARS		15 YEARS AND OVER		AS % OF TOTAL CURRENT EXPENDITURE ON EDUCATION	PER PUPIL AS % OF GNP PER CAPITA							1997	1997	
	BOYS	GIRLS	BOYS	GIRLS	BOYS	GIRLS	BOYS	GIRLS	MALE	FEMALE	MALE	FEMALE											
ANTIGUA/ BARBUDA															7 380			91	463	36			ANTIGUA/ BARBUDA
ARGENTINA							6	4	98	99	97	96	46	8	8 950		4	123	223	88	3	73	ARGENTINA
BAHAMAS	96	97							97	99	95	96					0	99	230	66	3	74	BAHAMAS
BARBADOS	78	77															0	199	285	47	2	76	BARBADOS
BELIZE							12	9					63	11	2 670		2		183	47	4	75	BELIZE
BOLIVIA	95	87							97	92	91	77	51	10	970	7	13	55	116	61	4	61	BOLIVIA
BRAZIL									90	93	84	84	54	10	4 790	29	15	40	223	78	2	67	BRAZIL
BRIT. VIRGIN ISLANDS																			218	56			BRIT. VIRGIN ISLANDS
CHILE	90	88	91	88	100	100	6	4	98	99	95	95	60	11	4 820	15	0	98	215	84	2	75	CHILE
COLOMBIA					70	76	8	7	96	97	91	91	41	9	2 180	7	6	46	115	73	3	70	COLOMBIA
COSTA RICA	86	87	89	89	86	89	13	10	98	98	95	95	40	13	2 680	19	5	88	140	49	3	76	COSTA RICA
CUBA	92	92	100	100					100	100	96	96	32	18			0	118	239	76	2	76	CUBA
DOMINICA															3 040				78	69			DOMINICA
DOMINICAN REPUBLIC									90	91	83	82	50	4	1 750	20	15	52	95	62	3	71	DOMINICAN REPUBLIC
ECUADOR					84	86	4	3	97	97	93	89	38	6	1 570	30	5	70	130	59	3	70	ECUADOR
EL SALVADOR			78	78	76	77	5	4	88	86	80	74	64	7	1 810		15	48	677	45	3	69	EL SALVADOR
GRENADA															3 140				353	36			GRENADA
GUATEMALA			76	69	52	47	16	14	84	70	74	59	63	6	1 580	53	15	33	61	39	5	64	GUATEMALA
GUYANA	93	93	87	87	91	92	5	3	100	100	99	97	71	10	800		0	50	55	35	2	64	GUYANA
HAITI	22	23							60	59	48	43			380		24	3	5	32	4	54	HAITI
HONDURAS									78	79	71	70	53	10	740	47	8	55	95	44	4	69	HONDURAS
JAMAICA	96	96							89	96	81	90	31	10	1 550	5	0	63	183	54	3	75	JAMAICA
MEXICO			100	100	85	86	8	6	97	95	92	88	50	12	3 700	15	6	97	272	73	3	72	MEXICO
NETHERLANDS ANTILLES									98	98	96	96						334	328	69	2	75	NETHERLANDS ANTILLES
NICARAGUA	71	73	76	79	43	52	16	14	67	68	63	63		13	410	44	13	30	68	62	4	68	NICARAGUA
PANAMA	91	92							97	96	92	90	30	9	3 080	26	3	62	187	56	3	74	PANAMA
PARAGUAY	94	92	91	91	77	80	10	8	97	97	94	91	50	10	2 000		7	43	101	52	4	70	PARAGUAY
PERU			91	90			16	15	98	94	94	84	35	5	2 610	49	2	84	126	71	3	68	PERU
ST. KITTS ET NEVIS													38	6	6 260				264	34			ST. KITTS ET NEVIS
ST. LUCIA													45	10	3 510				213	37			ST. LUCIA
ST. VINCENT/ GRENADINES															2 420			9	163	48			ST. VINCENT/ GRENADINES
SURINAME											95	92			1 320		0	122	153	49	2	70	SURINAME
TRINIDAD ET TOBAGO	91	91	88	88	98	97	6	5	100	100	99	97	41	10	4 250		0	123	333	72	2	74	TRINIDAD ET TOBAGO
URUGUAY	91	92	92	93	96	99	11	8	99	100	97	98	33	8	6 130		2	293	239	90	2	74	URUGUAY
VENEZUELA	87	89	83	85	86	92	12	9	97	98	93	92			3 480	12	1	206	180	86	3	72	VENEZUELA

BEST COPY AVAILABLE

ARAB STATES

COUNTRIES/ TERRITORIES	NET ENROLMENT RATIO				PERCENTAGE OF 1995 COHORT REACHING GRADE 5		PERCENTAGE OF REPEATERS 1996		ESTIMATED ADULT LITERACY RATES 1997				PUBLIC CURRENT EXPENDITURES IN PRE-PRIMARY AND PRIMARY EDUCATION[1]		GNP PER CAPITA - US $ 1997	% PEOPLE LIVING ON LESS THAN $1 A DAY (PPP) 1981-1995	ESTIMATED % OF ECONOMICALLY ACTIVE CHILDREN (10-14 YEARS)	DAILY NEWSPAPER PER 1,000 INHABITANTS 1996	TELEVISION RECEIVERS 1,000 INHABITANTS 1997	% URBAN POPULATION 1996	TOTAL FERTILITY RATE	LIFE EXPECTANCY AT BIRTH	COUNTRIES/ TERRITORIES
	1990		1996						15-24 YEARS		15 YEARS AND OVER		AS % OF TOTAL CURRENT EXPENDITURE ON EDUCATION	PER PUPIL AS % OF GNP PER CAPITA							1997	1997	
	BOYS	GIRLS	BOYS	GIRLS	BOYS	GIRLS	BOYS	GIRLS	MALE	FEMALE	MALE	FEMALE											
ALGERIA	99	87	97	91	93	95	13	8	86	68	73	48			1 500	2	1	38	105	56	3,8	69	ALGERIA
BAHRAIN	99	99	97	99	94	95	5	5	98	98	90	81	28	8			0	117	472	90	2,9	73	BAHRAIN
DJIBOUTI	37	27	36	27					100	100	62	35							45	82	5,3	50	DJIBOUTI
EGYPT			98	88			7	5	75	59	65	40	**67**	**12**	1 200	8	10	38	119	45	3,4	66	EGYPT
IRAQ			81	71													3	20	83	75	5,3	62	IRAQ
JORDAN							1	1	97	100	92	82	**65**	**19**	1 520	3	0	42	82	71	4,9	70	JORDAN
KUWAIT	45	44	62	61	94	97	3	3	91	92	83	77	**69**	**14**			0	377	505	97	2,9	76	KUWAIT
LEBANON									97	92	91	78	**69**	**6**	3 350		0	141	375	88	2,7	70	LEBANON
LIBYAN ARAB JAMAHIRIYA									100	90	89	63					0	14	140	85	3,8	70	LIBYAN ARAB JAMAHIRIYA
MAURITANIA			61	53	61	68	15	17	56	38	49	28	39	11	440	31	23	1	25	51	5,5	54	MAURITANIA
MOROCCO	68	48	83	65	76	74	14	10	74	53	59	33	35	11	1 260	1	4	27	115	52	3,1	67	MOROCCO
OMAN	73	68	70	68	96	96	11	7	99	92	77	55	42				0	28	694	76	5,9	71	OMAN
PALESTINIAN AUTONOMOUS TERRITORIES					99	100	4	3												94	7,3	71	PALESTINIAN AUTONOMOUS TERRITORIES
QATAR	87	86			97	99	6	3	92	96	80	81					0	161	404	91	3,7	72	QATAR
SAUDI ARABIA	65	53	63	60	87	92	11	4	95	88	82	63	**82**	**19**	7 150		0	59	262	83	5,8	71	SAUDI ARABIA
SOMALIA																	32	1	15	26	7,3	47	SOMALIA
SUDAN					75	73			81	66	65	41			290		29	27	86	31	4,6	55	SUDAN
SYRIAN ARAB REPUBLIC	100	93	95	87	93	94	8	6	94	75	87	57	42	8	1 120		4	20	70	52	4,0	69	SYRIAN ARAB REPUBLIC
TUNISIA	97	90	99	96	90	92	17	15	96	85	78	56	43	15	2 110	4	0	31	100	62	2,6	70	TUNISIA
UNITED ARAB EMIRATES	95	93	79	78	83	84	5	3	86	93	74	77					0	170	134	84	3,4	75	UNITED ARAB EMIRATES
YEMEN									81	39	64	21			270		20	15	29	34	7,6	58	YEMEN

1. Bold figures include secondary education.

EASTERN ASIA/OCEANIA

COUNTRIES/ TERRITORIES	NET ENROLMENT RATIO				PERCENTAGE OF 1995 COHORT REACHING GRADE 5		PERCENTAGE OF REPEATERS 1996		ESTIMATED ADULT LITERACY RATES 1997				PUBLIC CURRENT EXPENDITURES IN PRE-PRIMARY AND PRIMARY EDUCATION[1]		GNP PER CAPITA - US $ 1997	% PEOPLE LIVING ON LESS THAN $1 A DAY (PPP) 1981-1995	ESTIMATED % OF ECONOMICALLY ACTIVE CHILDREN (10-14 YEARS)	DAILY NEWSPAPER PER 1,000 INHABITANTS 1996	TELEVISION RECEIVERS 1,000 INHABITANTS 1997	% URBAN POPULATION 1996	TOTAL FERTILITY RATE	LIFE EXPECTANCY AT BIRTH	COUNTRIES/ TERRITORIES
	1990		1996						15-24 YEARS		15 YEARS AND OVER		AS % OF TOTAL CURRENT EXPENDITURE ON EDUCATION	PER PUPIL AS % OF GNP PER CAPITA							1997	1997	
	BOYS	GIRLS	BOYS	GIRLS	BOYS	GIRLS	BOYS	GIRLS	MALE	FEMALE	MALE	FEMALE											
BRUNEI DARUSSALAM	92	90	90	91					99	100	94	86											
CAMBODIA					51	46	27	25									0	69	250	69	3	76	BRUNEI DARUSSALAM
CHINA	99	95	100	100	93	94	2	1	99	95	91	75	37	6	300		24	2	9	20	5	53	CAMBODIA
CHINA, HONG KONG SAR			88	91					99	100	96	88	22	6	860	29	10	45	321	30	2	70	CHINA
COOK ISLANDS															25 200			786	283	95	1	78	CHINA, HONG KONG SAR
D.P.R. KOREA													61					105	193	60			COOK ISLANDS
FIJI	100	100							99	99	94	89					0	199	52	61	2	72	D.P.R. KOREA
INDONESIA	100	95	96	93	88	89	6	6	98	96	91	80	**74**		2 460		0	51	27	41	3	73	FIJI
KIRIBATI					95	93	0	0							1 110	15	9	23	68	35	3	65	INDONESIA
LAO PEOPLE'S DEM. REP.	66	57	76	68	57	54	25	22					55	7	910				15	36			KIRIBATI
MACAU	81	81							99	96	96	89			400		26	4	10	21	6	53	LAO PEOPLE'S DEM. REP.
MALAYSIA			100	100					97	97	90	81	37	10	4 530	6	3	455	109	99	1	78	MACAU
MONGOLIA			79	83	89	92	1	1	99	100	99	99	20	11	390		2	163	172	54	3	72	MALAYSIA
MYANMAR									91	89	89	79	48	3			24	27	47	61	3	66	MONGOLIA
PAPUA NEW GUINEA									100	100	82	65			930		18	10	5,9	26	2	60	MYANMAR
PHILIPPINES									98	98	95	94	55	9	1 200	28	7	15	9,3	16	5	59	PAPUA NEW GUINEA
REPUBLIC OF KOREA	100	100	92	93	98	99							45	11	10 550		0	82	52	54	4	68	PHILIPPINES
SAMOA			96	95											1 140			20	348	81	2	72	REPUBLIC OF KOREA
SINGAPORE			93	92					99	100	96	87	26	7	32 810		0	85	61	21	4	71	SAMOA
SOLOMON ISLANDS										100		100			870		27	324	388	100	2	77	SINGAPORE
THAILAND									99	98	97	93	50	14	2 740		15	64	6,2	17	5	72	SOLOMON ISLANDS
TONGA															1 810			72	254	20	2	69	THAILAND
TUVALU																			21	41			TONGA
VANUATU													58	14	1 340					47			TUVALU
VIET NAM									97	97	95	89	43	7	310		8	4	14 47	19 19	4 3	67 67	VANUATU VIET NAM

1. Bold figures include secondary education.

BEST COPY AVAILABLE

SOUTHERN ASIA

COUNTRIES	NET ENROLMENT RATIO				PERCENTAGE OF 1995 COHORT REACHING GRADE 5		PERCENTAGE OF REPEATERS 1996		ESTIMATED ADULT LITERACY RATES 1997				PUBLIC CURRENT EXPENDITURES IN PRE-PRIMARY AND PRIMARY EDUCATION[1]		GNP PER CAPITA US $ 1997	% PEOPLE LIVING ON LESS THAN $1 A DAY (PPP) 1981-1995	ESTIMATED % OF ECONOMICALLY ACTIVE CHILDREN (10-14 YEARS)	DAILY NEWSPAPER PER 1,000 INHABITANTS 1996	TELEVISION RECEIVERS 1,000 INHABITANTS 1997	% URBAN POPULATION 1996	TOTAL FERTILITY RATE	LIFE EXPECTANCY AT BIRTH	COUNTRIES
	1990		1996						15-24 YEARS		15 YEARS AND OVER		AS % OF TOTAL CURRENT EXPENDITURE ON EDUCATION	PER PUPIL AS % OF GNP PER CAPITA							1997	1997	
	BOYS	GIRLS	BOYS	GIRLS	BOYS	GIRLS	BOYS	GIRLS	MALE	FEMALE	MALE	FEMALE											
AFGHANISTAN									65	34	48	18					25	6	13	20	7	45	AFGHANISTAN
BANGLADESH	68	60							58	37	50	27	45	5	360		29	9	6	18	3	58	BANGLADESH
BHUTAN							20	18			58	30	44	30	430		53		6	6	6	61	BHUTAN
INDIA							3	4	77	56	67	39	40	11	370	53	13	35	65	27	3	63	INDIA
IRAN, ISLAMIC REPUBLIC	100	96	91	88			7	4					29	6	1 780		4	26	71	59	3	69	IRAN, ISLAMIC REPUBLIC
MALDIVES							7		98	99	96	96			1 180		5	19	28	27	5	65	MALDIVES
NEPAL									73	37	56	21	49	10	220	53	44	11	6	10	4	57	NEPAL
PAKISTAN									69	39	55	25	48	9	500	12	17	21	22	34	5	64	PAKISTAN
SRI LANKA					83	84	3	2	97	96	94	88	**7.5**	**9**	800	4	2	29	84	22	2	73	SRI LANKA

1. Bold figures include secondary education.

Main data sources used in preparing the tables and graphics

■ Monitoring Learning Achievement (MLA)

Vinayagum Chinapah et al., With Africa for Africa: Towards Quality Education for All, Draft Regional Report, EFA 2000 Assessment, MLA Project, 1999.

Vinayagum Chinapah, Handbook on Monitoring Learning Achievement: Towards Capacity building, UNESCO-UNICEF Follow-up to Jomtien, UNESCO, Paris, 1997.

■ Southern Africa Consortium for Measuring Educational Quality (SACMEQ)

Kulpoo Dhurumbeer, "*The Quality of Education: Some Policy Suggestions Based on a Survey of Schools, Mauritius*", International Institute for Educational Planning, Ministry of Education and Human Resource Development, Mauritius, SACMEQ Policy Research: Report No. 1 (Series Editor : Kenneth Ross), UNESCO, Paris, 1998.

Voigts Friedhelm, "*The Quality of Education: Some Policy Suggestions Based on a Survey of Schools, Namibia,*" International Institute for Educational Planning, Ministry of Basic Education and Culture, Namibia, SACMEQ Policy Research: Report No. 2 (Series Editor: Kenneth Ross), UNESCO, Paris, 1998.

Thomas Machingaidze, Patrick Pfukani and Sibangani Shumba, "*The Quality of Education: Some Policy Suggestions Based on a Survey of Schools, Zimbabwe*", International Institute for Educational Planning, Ministry of Education and Culture, Zimbabwe. SACMEQ Policy Research: Report No. 3 (Series Editor : Kenneth Ross), UNESCO, Paris, 1998.

Sebtuu Nassor and Khadija Ali Mohammed, "*The Quality of Education: Some Policy Suggestions Based on a Survey of Schools, Tanzania (Zanzibar)*", International Institute for Educational Planning, Ministry of Education and Human Resource Development, Tanzania, SACMEQ Policy Research: Report No. 4 (Series Editor : Kenneth Ross), UNESCO, Paris, 1998.

Manasseh Nkamba and Joe Kanyika, "*The Quality of Education: Some Policy Suggestions Based on a Survey of Schools, Zambia*", International Institute for Educational Planning, Ministry of Education, Zambia, SACMEQ Policy Research: Report No. 5 (Series Editor : Kenneth Ross), UNESCO, Paris, 1998.

■ Latin American Laboratory for Assessment of Educational Quality (Laboratorio)

Latin American Educational Quality Assessment Laboratory. First International Comparative Study of Language, Mathematics, and Associated Factors in Third and Fourth Grades, UNESCO, 1998.

■ Third International Mathematics and Science Study (TIMSS)

Michael O. Martin and Dana L. Kelly, TIMSS Technical Reports, Implementation Analysis Vol. 2, Boston College, Chestnut Hill, Massachusetts, USA, 1997.

TIMSS International Database (in CD Rom), Primary School Years - Population 1, 1995 Assessment, 1997.

TIMSS International Database (in CD Rom), Middle School Years - Population 2, 1995 Assessment, 1997.

■ Programme for the Analysis of Educational Systems of the CONFEMEN Countries (PASEC)

Programme d'Analyse des Systemes Educatifs (PASEC), Les facteurs de l'efficacite de l'enseignement primaire: donnees et resultats sur cinq pays d'Afrique et de l'Ocean Indien, CD-ROM (Version de mars 1999).

■ United Nations Development Programme (UNDP)

Human Development Report 1997, New York, Oxford University Press, 1997.

Human Development Report 1998, New York, Oxford University Press, 1998.

■ UNESCO

Adult Education in a Polarizing World, Education for All: Status and Trends 1997, Secretariat of EFA Forum, Paris, 1997.

Wasted Opportunities, When schools fail: Status and Trends 1998, Secretariat of EFA Forum, Paris, 1998.

Compendium of Statistics on Illiteracy, UNESCO, Division of Statistics, Paris, 1995.

Statistical Yearbook, UNESCO Institute for Statistics, UNESCO Publishing, Paris, 1980-98.

World Education Indicators Database on Internet, www.unescostat.unesco.org

World Education Report 1999, UNESCO Publishing, Paris, 1998.

■ UNICEF

State of the World's Children 1998, New York, UNICEF, 1999.

■ United Nations

World Population Prospects: The 1998 Revision, New York, 1998.

■ World Bank

World Development Report 1998, Washington D.C., World Bank, 1998.

How much are our children learning at school? Is it enough? Are there major gaps in their learning that need to be addressed? For tens of millions of parents around the world, these questions are deeply troubling. They and their children live in developing countries or in marginalized areas of the industrialized world where schools are underfunded, teachers lack adequate training, and student achievement may or may not reach the level of basic literacy and numeracy.

A number of research projects launched in the wake of the World Conference on Education for All, which met in Jomtien, Thailand in 1990, have documented the extent to which knowledge, like other forms of wealth, seems to be unevenly distributed in the world of the 21st century. These studies have been carried out on international, regional and national bases, and they have documented disparities in learning achievement not only between the various regions and countries of the world but within particular nations.

This fifth issue of Education for All: Status and Trends examines the reasons for the growing global interest in educational measurement, lays out some of the central findings of the movement and then takes up the all-important question of how assessment of student outcomes can be put to the service of promoting quality education for all children.

This report is published by UNESCO for the International Consultative Forum on Education for All, the global mechanism established to promote and monitor progress toward the goal of Education for All (EFA).

For further information, please contact:
EFA Forum Secretariat

UNESCO

7, Place de Fontenoy
75352 Paris 07 SP, France

fax: +33 (0)1 45 68 56 29

e-mail: efa@unesco.org

Internet web site:
www.education.unesco.org/efa

BEST COPY AVAILABLE